

A Metodologia de DNA *Shuffling* na Produção de Diversidade Gênica



República Federativa do Brasil

Luiz Inácio Lula da Silva

Presidente

Ministério da Agricultura, Pecuária e Abastecimento

Roberto Rodrigues

Ministro

Empresa Brasileira de Pesquisa Agropecuária

Conselho de Administração

José Amauri Dimázio

Presidente

Clayton Campanhola

Vice-Presidente

Alexandre Kalil Pires

Dietrich Gerhard Quast

Sérgio Fausto

Urbano Campos Ribeiral

Membros

Diretoria-Executiva da Embrapa

Clayton Campanhola

Diretor-Presidente

Gustavo Kauark Chianca

Herbert Cavalcante de Lima

Mariza Marilena T. Luz Barbosa

Diretores-Executivos

Embrapa Recursos Genéticos e Biotecnologia

Luiz Antonio Barreto de Castro

Chefe -Geral

Clara Oliveira Goedert

Chefe-Adjunto de Pesquisa e Desenvolvimento

José Manuel Cabral de Sousa Dias

Chefe-adjunto de Comunicação e Negócios

Arthur da Silva Mariante

Chefe-Adjunto de Administração



ISSN 0102 - 0110

Dezembro, 2003

Empresa Brasileira de Pesquisa Agropecuária
Centro Nacional de Pesquisa de Recursos Genéticos e Biotecnologia
Ministério da Agricultura, Pecuária e Abastecimento

Documentos 108

A metodologia de DNA *shuffling* na produção de diversidade gênica

Maria Cristina Mattar da Silva
Edson Luiz Zangrando Figueira
Maria Fátima Grossi de Sá

Brasília, DF
2003

Exemplares desta publicação podem ser adquiridos na:

Embrapa - Recursos Genéticos e Biotecnologia

Serviço de Atendimento ao Cidadão
Parque Estação Biológica, Av. W5 Norte (Final) - Brasília, DF
CEP 70770-900 - Caixa Postal 02372
PABX: (61) 448-4600
Fax: (61) 340-3624
<http://www.cenargen.embrapa.br>
e.mail:sac@cenargen.embrapa.br

Comitê de Publicações da Unidade

Presidente: José Manuel Cabral de Sousa Dias
Secretária-Executiva: Maria José de Oliveira Duarte
Membros: Maurício Machaim Franco
Regina Maria Dechechi G. Carneiro
Luciano Lourenço Nass
Sueli Correa Marques de Mello
Vera Tavares Campos Carneiro

Suplentes: Maria Alice Bianchi
Maria Fátima Batista

Supervisor Editorial: Maria José de Oliveira Duarte
Normalização Bibliográfica: Maria Alice Bianchi
Tratamento de Ilustrações: Altevir de Carvalho Freitas
Editoração Eletrônica: Altevir de Carvalho Freitas
Capa: Altevir de Carvalho Freitas

1ª edição

1ª impressão (2003): tiragem 150

Todos os direitos reservados.

A reprodução não autorizada desta publicação, no todo ou em parte, constitui violação dos direitos autorais (Lei nº 9.610).

Silva, Maria Cristina Mattar da.

A metodologia de DNA *shufflig* na produção de diversidade gênica / Maria Cristina Mattar da Silva, Edson Luiz Zangrando Figueira, Maria Fátima Grossi de Sá. — Brasília : Embrapa Recursos Genéticos e Biotecnologia, 2003.

30 p. — (Documentos / Embrapa Recursos Genéticos e Biotecnologia, ISSN 0102-0110 ; n. 108)

1. DNA *shuffling* – Metodologia. 2. DNA *shuffling* – Biotecnologia. 3. DNA *shuffling* – Melhoramento genético vegetal. I. Silva, Maria Cristina Mattar da. II. Figueira, Edson Luiz Zangrando. III. Sá, Maria Fátima Grossi de. IV. Título. V. Série.

576.5 - CDD 21

Sumário

Resumo	06
1. Introdução	07
2. A metodologia de <i>DNA shuffling</i>	08
3. Modificações do método original de <i>DNA shuffling</i>	10
3.1 Método de <i>Family shuffling</i>	10
3.1.1 <i>DNA shuffling</i> utilizando fita simples	10
3.1.2 <i>DNA shuffling</i> utilizando fita dupla e enzimas de restrição	11
3.2 Método SPOP (“ <i>Shuffled Proteins On Phages</i> ”)	12
3.3 Método SCRATHY	12
3.4 <i>DNA shuffling</i> sintético e DHR (<i>Degenerate Homoduplex Recombination</i>)	13
4. Métodos computacionais para prever recombinações.	15
5. Aplicações da metodologia de <i>DNA shuffling</i> no melhoramento genético de plantas	16
5.1 Genes marcadores	16
5.2 Tolerância a herbicidas	16
5.3 Resistência a insetos	17
5.3.1 Biblioteca combinatória de proteínas Cry	17
5.3.2 Biblioteca combinatória de inibidores de α -amilases	18
7. Referências bibliográficas	19

O método de DNA *shuffling* (“Embaralhamento de DNA”) é empregado para recombinar seqüências homólogas de DNA randomicamente fragmentadas, durante a evolução molecular *in vitro*. O procedimento original é constituído por diferentes etapas, incluindo a obtenção de substrato (genes parentais a serem recombinados), digestão enzimática dos genes parentais, recombinação e recuperação de novas seqüências, através de repetidos ciclos de anelamento na presença da enzima DNA polimerase. O método inicialmente proposto, denominado *Family shuffling* sofreu modificações a fim de minimizar as limitações relacionadas à baixa taxa de recombinação entre seqüências que possuem alta homologia. Em todas as situações, obtêm-se infinitas variantes que são selecionadas segundo características previamente estabelecidas. A metodologia apresenta-se promissora com perspectivas de aplicação nos processos biotecnológicos, em especial no melhoramento genético de plantas, a fim de desenvolver novos genes que conferem tolerância a herbicidas e resistência a insetos-praga ou doenças. Neste contexto, nossa equipe, na Embrapa Recursos Genéticos e Biotecnologia, vem aplicando a metodologia do DNA *shuffling* para gerar bibliotecas de genes mutantes de inibidores de α -amilases e toxinas Cry de *Bacillus thuringiensis*. As moléculas com atividades novas ou melhoradas, selecionadas dessas bibliotecas, serão utilizadas no controle de pragas de importância econômica na agricultura.

Durante a última década, vários laboratórios exploraram a possibilidade de reproduzir processos evolucionários, *in vitro*, com o intuito de desenvolver proteínas com características intrínsecas melhoradas. A exemplo, disponibilizaram-se enzimas com novas propriedades, maior atividade e especificidade, através de procedimentos envolvendo ciclos sucessivos de mutagênese *in vitro* (Caldwell & Joyce, 1992; Chen & Arnold, 1993; You & Arnold, 1994; Reidhaar-Olson et al., 1994). No entanto, o melhoramento das propriedades das proteínas requer grandes populações de mutantes em associação a processos de seleção estridente, devido à alta taxa de mutações randômicas deletérias ou neutras. Frente ao problema, Stemmer (1994 a,b) inovou com a introdução da tecnologia do *DNA shuffling* (“embaralhamento de DNA”), proporcionando melhorias nas técnicas de evolução molecular dirigida de modo a favorecer mutações desejáveis e remover as deletérias.

A tecnologia do *DNA shuffling* vem sendo disseminada com sucesso no meio científico, como verificado recentemente no “7th International Congress of Plant Molecular Biology” em Barcelona (Lassner, 2003). Trata-se de uma ferramenta poderosa para a construção de bibliotecas combinatórias de genes mutantes, cujo produto resulta na geração de grande variedade de novos genes e permite a seleção das proteínas com características almejadas, segundo critérios previamente estabelecidos. Em nível comercial, as enzimas são as mais susceptíveis de aplicação da tecnologia, minimizando os problemas de incompatibilidade entre os processos industriais e a atividade enzimática em condições drásticas de valores de pH, altas temperaturas e força iônica (Nedwin, 1997). O método, ainda permite desenvolver compostos inócuos aos humanos, aumentando a biossegurança alimentar e gerando novas proteínas distintas das existentes e patenteadas. O aprimoramento dos protocolos de *DNA shuffling* tem permitido estudos envolvendo análises estruturais da interação proteína-proteína, resultando em excelente alternativa nas situações em que as moléculas não possuem estrutura cristalina determinada.

Esta revisão tem como objetivo descrever as etapas do método original de *DNA shuffling* e apresentar as estratégias para minimizar as limitações do protocolo original. Em adição, encontram-se exemplos de aplicações, desta tecnologia, no melhoramento de plantas e perspectivas de utilização na obtenção de bibliotecas de genes mutantes, codificando proteínas melhoradas envolvidas no controle de insetos-praga de interesse econômico na agricultura.

2. A metodologia de *DNA shuffling*

A metodologia do *DNA shuffling* foi desenvolvida para estudar a evolução molecular dos genes codificadores de β -lactamase (Stemmer, 1994a, 1994b) e, posteriormente, aplicada com o intuito de obter proteínas funcionalmente melhoradas (Zhao&Arnold, 1997).

O protocolo original envolve as seguintes etapas: (i) isolamento do gene ou dos genes a serem fragmentados; (ii) digestão enzimática do(s) gene(s), (iii) recombinação dos fragmentos gênicos e (iv) amplificação das novas seqüências. A Figura 1 apresenta a ilustração para estas etapas, detalhadas a seguir:

(1) Obtenção do substrato. O substrato da reação consiste de gene(s) apresentando variada taxa de homologia, isolados através de digestão com enzimas de restrição ou amplificação em Reação de Polimerase em Cadeia (PCR). Os oligonucleotídeos iniciadores, utilizados na obtenção do gene, devem conter sítios para enzimas de restrições a fim de permitir futuras subclonagens em vetores de expressão.

(2) Fragmentação do(s) gene(s). O substrato é fragmentado aleatoriamente, utilizando principalmente a enzima *DNAse I* sob condições controladas. Alternativamente, utilizam-se enzimas de restrição, podendo neste caso, se estimar o tamanho dos fragmentos gerados (Kikuchi et al., 1999). Nesta etapa, é recomendável a obtenção de uma mistura de fragmentos contendo entre 50 a 300 pares de bases.

(3) Recombinação gênica. Uma PCR, sem adição de oligonucleotídeos iniciadores, é realizada utilizando como molde o produto da etapa (2). Os fragmentos isolados ao se anelarem formam “homoduplexes” (originados de anelamento de seqüências do mesmo gene) ou “heteroduplexes” (anelamento de seqüências de genes distintos) que funcionarão como “iniciadores” da reação. Uma taxa maior de heteroduplexes, em relação aos homoduplexes, é desejada com intuito de atingir maior diversidade na recombinação. Durante a etapa de extensão, os pequenos duplexes se sobrepõem, formando pontos de junção entre os diferentes segmentos de seqüências homólogas.

(4) Amplificação dos genes recombinados: Uma alíquota do produto da primeira PCR (etapa 3) é submetida à nova PCR, porém com adição de oligonucleotídeos iniciadores contendo seqüências 5' e 3' dos genes originais. Os homo ou heteroduplexes estendidos atuam como moldes para amplificação dos genes recombinados com o mesmo tamanho do gene original.

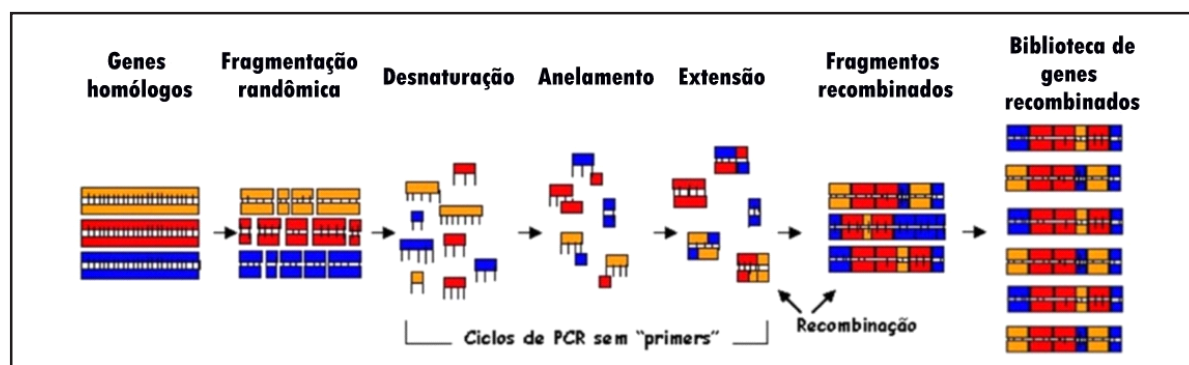


Fig. 1. Representação esquemática das etapas do método de *DNA shuffling*, partindo de uma família de genes homólogos e utilizando digestão enzimática para gerar fragmentação randômica. Os fragmentos hibridizam entre si durante a PCR, sem adição de oligonucleotídeos. A mistura de fragmentos hibridizados é submetida à outra PCR, utilizando oligonucleotídeos específicos, para obter a população de genes mutantes no tamanho do(s) gene (s) parentais, originando a biblioteca de recombinantes. (Adaptada de Moore & Maranas, 2002).

A população de genes recombinados é inserida em vetores de expressão apropriados para a seleção de proteínas recombinantes com características desejadas. Neste contexto, a combinação da metodologia de *DNA shuffling* com a técnica de *Phage Display* (Figura 2) tem sido utilizada para explorar a grande diversidade de mutantes, apresentados em bibliotecas de fagos, expressando proteínas funcionais. A progênie, proveniente do primeiro ciclo de *DNA shuffling*, poderá servir de genes parentais (substrato) para os ciclos subseqüentes, de modo que o processo poderá ser repetido até se atingir as propriedades desejadas.

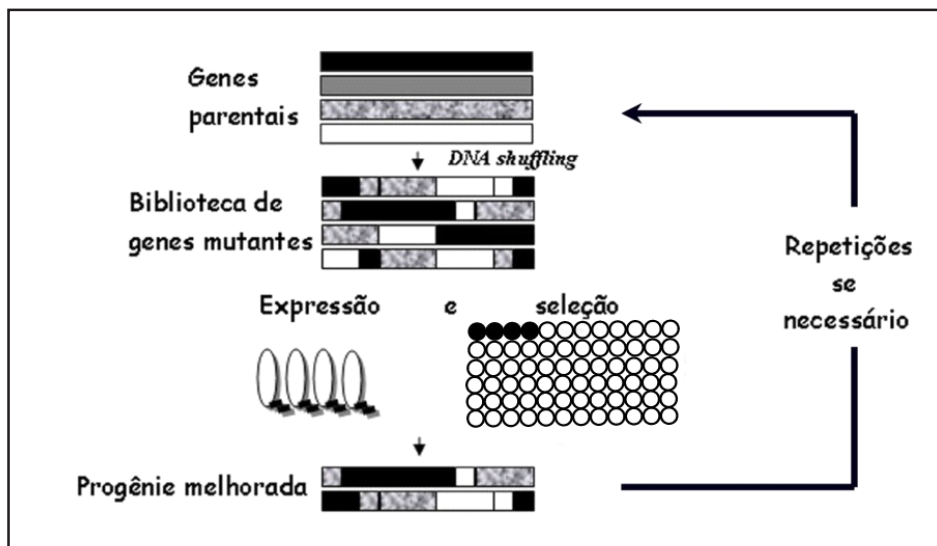


Fig. 2. Esquema de evolução molecular *in vitro* utilizando a técnica de *DNA shuffling* associada à metodologia de *Phage display*. A progênie resultante de um ciclo de *shuffling* pode ser utilizada como "genes parentais" para ciclos subseqüentes de hibridização, até a obtenção de mutantes com atividade almejada.

3. Modificações do método original de *DNA shuffling*

Embora bastante disseminado em experimentos de evolução molecular dirigida, o método original de *DNA shuffling* (Stemmer 1994a,b; Zhao&Arnold, 1997) tem sofrido diversas adaptações com o objetivo de superar limitações e aumentar o grau de diversidade de genes mutantes gerados.

Na tabela 1, encontra-se resumido as diferentes estratégias envolvendo o método de *DNA shuffling*.

Tabela 1. Sumário das estratégias envolvendo a metodologia de *DNA shuffling*

Métodos	Substratos para serem recombinados	Fragmentação	Benefícios e/ou desvantagens dos métodos	Referências
1. Family “shuffling”	1. A. DNA fita simples de genes homólogos e complementares entre si	1.A.1. Enzima <i>DNase I</i>	1.A.1.1. Apresenta redução na taxa de formação de homoduplexes.	Kikuchi et al., 2000
	1. B. DNA fita dupla de dois ou mais genes homólogos	1.B.1. Enzima <i>DNase I</i>	1.B.1.1. Pode gerar maior número de mutantes funcionais; 1.B.1.2. Apresenta alta taxa de formação de homoduplexes.	Crameri et al., 1998
		1.B.2. Enzimas de restrição	1.B.2.1. Possibilita a predição do número e o tamanho dos fragmentos gênicos; 1.B.2.2. A diversidade de mutantes fica limitada à existência de sítios de restrição; 1.B.2.3. Pode elevar a taxa de recombinação gênica e propiciar maior diversidade.	Kikuchi et al., 1999
2. SPOP (“Shuffled Proteins on Phages”)	2.A. cDNA amplificado por PCR	2.A.1. Enzima <i>DNase I</i>	2.A.1.1. Envolve ciclos consecutivos de <i>DNA shuffling</i> , confecção de biblioteca combinatoria em fagos e seleção funcional. 2.A.1.2. Pode gerar mutantes interessantes para os estudos de interação proteína-proteína	Stoop et al., 2000
3. Scratthy (Itchy + <i>DNA shuffling</i>)	3.A. Biblioteca de fragmentos de genes fusionados, truncados e selecionados de acordo com a função.	3.A.1. Enzima <i>DNase I</i>	3.A.1.1. Possibilita recombinações gênicas independentes da homologia de seqüências.	Lutz et al., 2001
4. <i>DNA shuffling</i> sintético e DHR (Recombinação de Homoduplexes Degenerados)	4.A. Fragmentos gênicos desenhados e sintetizados contendo todos os polimorfismos existentes entre seqüências gênicas de espécies diferentes.	Não se aplica	4.A.1. Aumenta a taxa de cruzamentos de alta resolução e a recombinação de genes independe da homologia de seqüências.	Ness et al., 2002; Coco et al., 2002

3.1 Método de *Family Shuffling*

Family Shuffling é a denominação para o método original descrito por Stemmer (1994b) quando da utilização de dois ou mais genes homólogos como substrato (Vieira & Messing, 1987; Zhang et al., 1997; Crameri et al., 1998; Yano et al., 1998; Kikuchi et al., 1999; Kikuchi et al., 2000). A baixa diversidade de mutantes encontrada em alguns experimentos de *shuffling* é, muitas vezes, atribuída à alta formação de homoduplexes oriunda de sobreposição dos fragmentos durante a etapa de recombinação utilizando PCR. Estratégias utilizando DNA fita dupla e enzimas de restrição, em substituição a enzima *DNase I* (Kikuchi et al., 1999) ou DNA fita simples como substrato (Kikuchi et al., 2000) são alternativas para amenizar este problema.

3.1.1 Utilizando DNA fita simples

A modificação no protocolo original consiste no preparo do substrato, cuja utilização de DNA fita simples reduz a formação de homoduplexes, permitindo o aumento da diversidade de mutantes (Figura 3). O substrato pode ser obtido a partir da clonagem do gene a ser recombinado em fagomídeo, que inicialmente é usado para transformar *Escherichia coli*, seguido por infecção com o fago auxiliar (M13K07), purificação das partículas virais contendo DNA fita simples e fragmentação com a enzima *DNase I*.

Este procedimento surgiu após aplicação da técnica de *Family shuffling* na transformação dos genes *xyl E* e *nah h*, ambos codificadores da enzima termoestável 2,3 dioxigenase catecol (C23Os), resultando em taxa de recombinação inferior a 1% (Kikuchi et al., 1999). O fato foi atribuído ao alto grau de identidade dos genes (80%), com predomínio na formação de homoduplexes em relação a heteroduplexes. Diante deste problema, a utilização de DNA fita simples favoreceu a formação de heteroduplexes e elevou o índice de recombinação para 14% (Kikuchi et al., 1999).

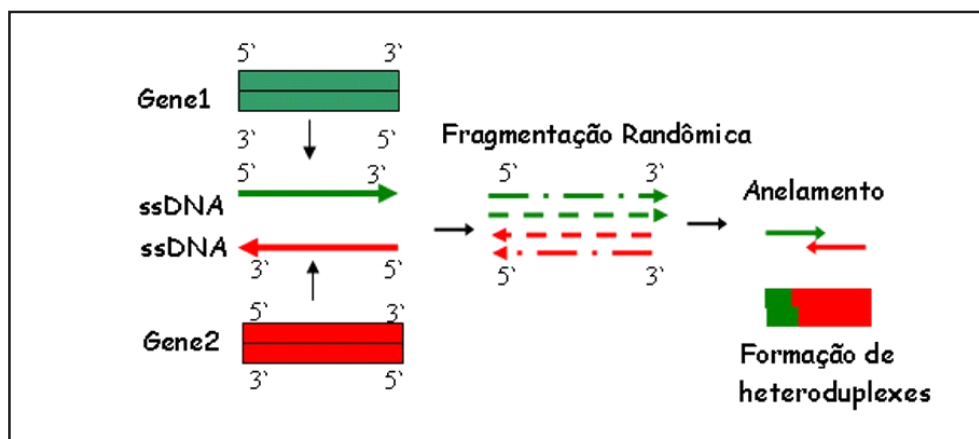


Fig. 3. Representação esquemática da técnica de *Family shuffling* utilizando fragmentos de DNAs de fita simples. Os DNAs fita simples e complementares entre si são purificados, a partir de partículas de fagos, e usados como substratos para o *shuffling*, favorecendo a formação de heteroduplexes e conseqüentemente, aumentando a diversidade de recombinantes. (Adaptada de Kikuchi-Nodate & Kagami, 2002)

3.1.2 Utilizando DNA fita dupla e enzimas de restrição

A fim de incrementar a taxa de recombinação, Kikuchi et al. (2000) sugeriram o uso de enzimas de restrições para a fragmentação de dois ou mais genes homólogos, em substituição a *DNase I* (Figura 4). Como aplicação, utilizaram novamente, os genes *xyl E* e *nah H* e atingiram taxa de 100% de recombinação.

Embora este procedimento apresente vantagens quanto à previsão do número e tamanho dos fragmentos gênicos, a diversidade dos produtos recombinados pode ficar restrita à disponibilidade de sítios de enzimas de restrições. Dependendo da seqüência gênica a digestão enzimática pode gerar fragmentos demasiadamente grandes, favorecendo à baixa taxa de recombinação. No entanto, a utilização de várias enzimas pode minimizar este problema e a decisão de utilizar esta alternativa, na obtenção de fragmentos de DNA, deve ser analisada em cada situação, considerando a análise minuciosa dos mapas de restrições dos genes homólogos.

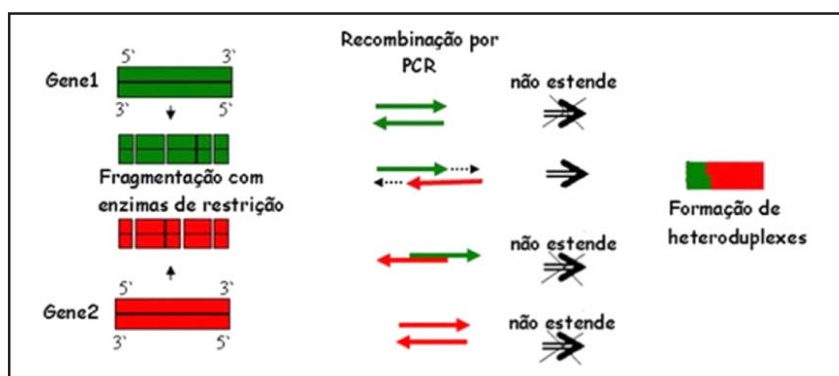


Fig. 4. Representação esquemática da técnica de *Family shuffling* utilizando fragmentos de DNAs de fita dupla obtidos com enzimas de restrição. O uso de enzima de restrição, ao invés de *DNase I*, durante a etapa de fragmentação impede que a sobreposição entre fragmentos do mesmo gene se estenda (dificultando a formação de homoduplexes). A elongação do DNA ocorre apenas nas moléculas hibridizadas, favorecendo o aumento da taxa de formação de heteroduplexes. (Adaptada de Kikuchi-Nodate & Kagami, 2002)

3.2 Método SPOP (“*Shuffled Proteins On Phages*”)

O método consiste de ciclos consecutivos de *DNA shuffling*, cada qual seguido da expressão da proteína alvo na superfície dos fagos (*Phage Display*) e posterior varredura para seleção de mutantes funcionais. Dois critérios são essenciais na aplicação do método SPOP: a proteína alvo deve (i) ser funcional após expressão em fagos e (ii) possuir características previamente estabelecidas para a seleção. Esta variação no método original de *DNA shuffling* permite estudos da relação estrutura-atividade ou interação entre proteínas, identificando os aminoácidos envolvidos no processo (Stoop et al., 2000).

Na validação do método, escolheu-se o sistema de ativação do inibidor I de plasminogênio humano (PAI-1) já que a molécula inibidora possui marca de seleção, estrutura tridimensional definida e pode servir de modelo nos estudos de estrutura-atividade de serino-proteinases e seus inibidores tipo serpinas. Após cinco ciclos de *DNA shuffling*, seguido de seleção mediada por anticorpos monoclonais e aplicação da técnica de ressonância plasmônica de superfície, elegeram-se aleatoriamente 27 clones, expressando mutantes de PAI-1, apresentando em média substituições de nove aminoácidos, distribuídas em 114 posições na superfície de interação da proteína. Muitos dos mutantes selecionados apresentaram apenas ligação parcial com alguns anticorpos monoclonais (clones defectivos). O alinhamento das seqüências desses clones (seleção negativa) possibilitou a identificação de um aminoácido dominante na ligação com cada um dos anticorpos monoclonais. Posteriormente, esses resultados foram confirmados através da construção de mutantes pontuais e avaliação da ligação com anticorpos monoclonais (Stoop et al., 2000).

3.3 Método SCRATHY

A estratégia denominada Scratchy (Lutz et al., 2001) surgiu a fim de evitar a recombinação tendenciosa entre genes em *loci* com alta homologia, em favorecimento aos de baixa, uma vez que a identidade entre as seqüências gênicas que compõem o substrato do método de *DNA shuffling* constitui uma limitação na diversificação de mutantes.

O procedimento envolve manipulação gênica adicional consistindo na combinação de dois métodos: o ITCHY (*Incremental Truncation for the Creation of Hybrid Enzymes*) e o *DNA shuffling*. O primeiro método consiste em gerar uma biblioteca de fusão entre os genes a serem utilizados como substrato no *DNA shuffling*. A combinação dos dois métodos é necessária, uma vez que a biblioteca de ITCHY permite que os membros que a compõem apresentem apenas uma recombinação por gene, e em combinação com o *DNA shuffling* torna-se possível introduzir recombinações múltiplas entre os genes de interesse.

Os genes que codificam para glicinamida ribonucleotideo-formiltransferase (GART) de *E. coli* (PurN) e de humano (hGART) (Figura 5) foram usados como modelos nos experimentos para validar a hipótese de que Scratchy pode gerar bibliotecas com recombinação múltipla entre genes de interesse, independente da homologia de seqüência (Lutz et al., 2001).

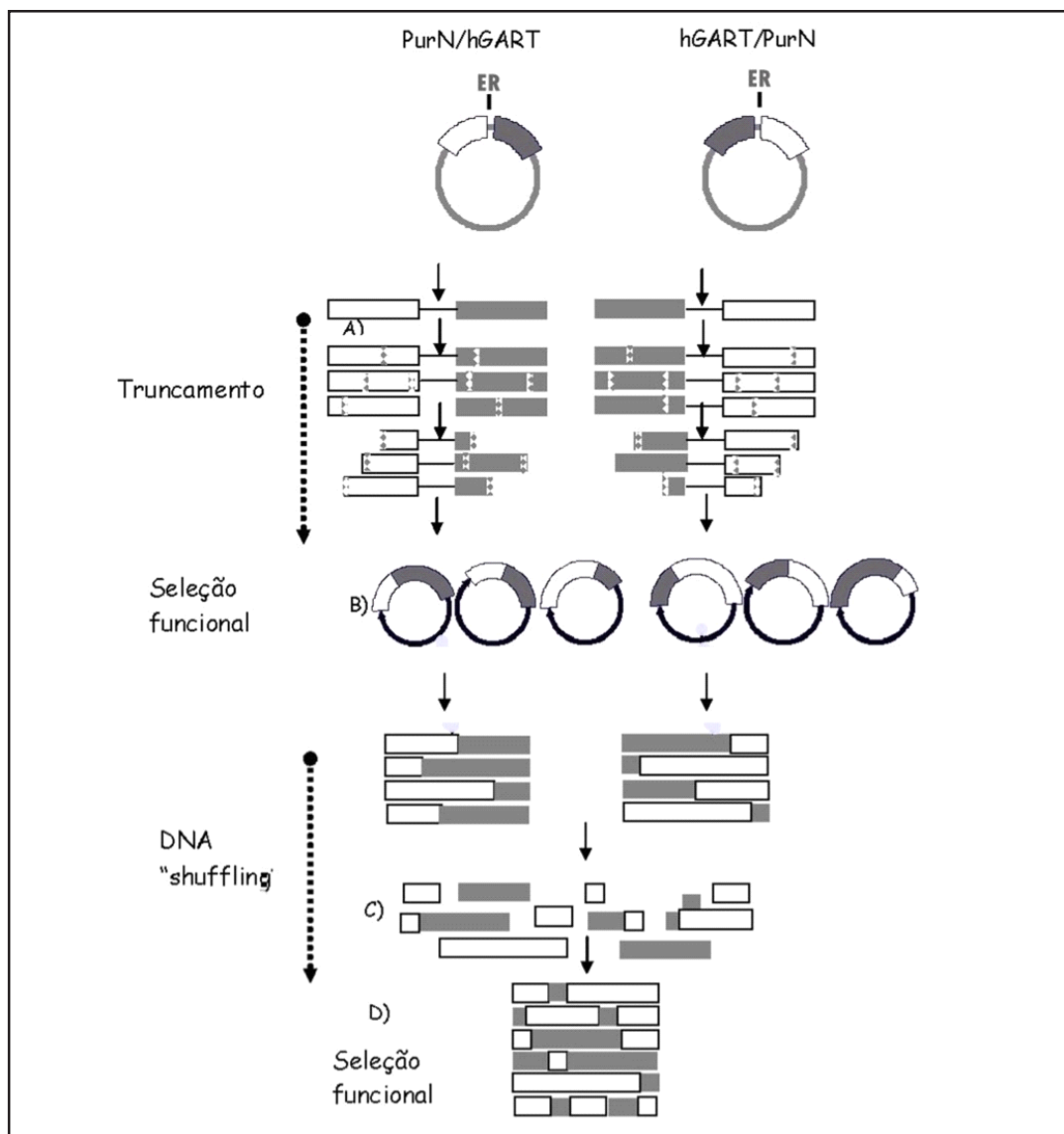


Fig. 5. Fluxograma do método Scratchy: (A) Biblioteca de fragmentos truncados (denominado Itchy) gerados a partir de construções fusionadas com os genes isolados de *E. coli* e humanos, codificadores para a enzima glicinamida ribonucleotideoformiltransferase, clonados em direções opostas. (B) Após a seleção funcional os transformantes são submetidos ao (C) DNA shuffling. (D) Uma seleção final identifica as construções que codificam para moléculas com atividades enzimáticas melhoradas. ER refere a sítios de enzimas de restrição. (Adaptado de Lutz et al., 2001).

3.4 DNA shuffling sintético e DHR (Recombinação de Homoduplexes Degenerados)

Recentemente, em alguns estudos de *shuffling* têm-se utilizado oligonucleotídeos sintéticos para suprir as limitações da metodologia original no que diz respeito à ausência de cruzamentos de alta resolução (devido, por exemplo, a sítios de recombinação muito próximos entre si) e a dependência na homologia dos genes (Ness et al., 2002; Coco et al., 2002; Ostermeier, 2003) (Figura 6).

Os métodos de recombinação usando oligonucleotídeos sintéticos são vantajosos sobre os outros métodos (Tabela 1), por possibilitar a obtenção de genes híbridos com cruzamentos de melhor resolução, devido à destruição de ligações entre pares de bases nas seqüências parentais, levando ao aumento da freqüência de cruzamentos em regiões de baixa homologia. Além disso, a utilização inicial de oligonucleotídeos permite a inserção de qualquer diversidade artificial desejada nos genes originais, possibilita a otimização com o uso de códons mais apropriados para cada microorganismo e permite o monitoramento da freqüência de um aminoácido em qualquer posição, a fim de aumentar a variabilidade da biblioteca.

A maioria dos métodos para DNA *shuffling in vitro* difere na forma como os fragmentos são gerados, mas recombina os genes usando amplificação por PCR. A estratégia denominada de DHR (Recombinação de Homoduplexes Degenerados) é diferente (Coco et al. 2002). Os oligonucleotídeos são desenhados e sintetizados para incorporar degenerações que envolvam todos os polimorfismos existentes entre os genes de diferentes espécies a serem recombinados. Os oligonucleotídeos da fita inferior (Figura 6) funcionam simplesmente como pontes, uma vez que eles não podem ser ligados por não possuírem o grupo 5' fosfato e não podem ser estendidos pela polimerase devido as modificações do 3' amino terminal. Esses oligonucleotídeos são usados como moldes para o restabelecimento do DNA fita simples superior, através do preenchimento dos espaços pela ação da enzima polimerase e pela união das extremidades. Até este ponto, o DHR não utiliza amplificação por PCR, pois o produto gênico consiste apenas de hibridização e ligação dos vários oligonucleotídeos, enquanto que o *shuffling* sintético envolve amplificação por PCR convencional. A partir de então, a diferença dos métodos está relacionada à temperatura. Enquanto o DHR utiliza acréscimo lento de temperatura na fase de anelamento (amplificação da fita superior), favorecendo a formação de homoduplexes, o *shuffling* sintético utiliza acréscimo rápido, o que favorece mais à formação de heteroduplexes e evita situações tendenciosas nas bibliotecas.

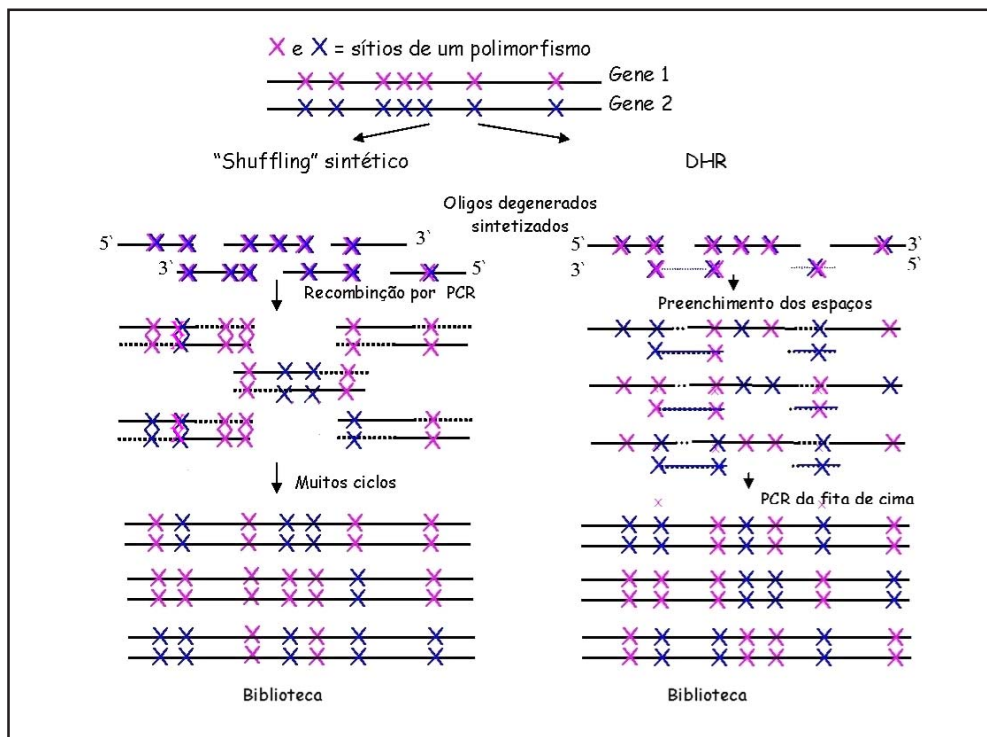


Fig. 6. Esquema comparativo dos métodos de DNA sintético e DHR (adaptado de Ostermeier, 2003) usando dois genes com sete pontos de polimorfismo ("x" em azul e cor-de-rosa). No DNA sintético o gene é dividido em comprimentos equivalentes, que se sobrepõem, formando oligonucleotídeos degenerados, que serão recombinados por PCR. No método DHR, em contraste, oligonucleotídeos são desenhados para minimizar a divergência na região de sobreposição e atuam como suportes para a recombinação que ocorre sem a utilização de PCR.

O método DHR é preferido, em muitas situações, por possibilitar a criação e avaliação de todas as combinações polimórficas. Contudo, a indução do aumento de diversidade é questionável (Ostermeier, 2003). Em muitos casos não é possível avaliar uma biblioteca completa porque a degeneração rapidamente excede o tamanho da biblioteca que pode ser construída ou avaliada. Portanto, apesar de se reconhecer que a diversidade é necessária para criar mutantes melhorados é preciso ter consciência que o aumento exagerado da diversidade de genes poderá ser contra-produtivo.

4. Métodos computacionais para prever recombinações

Os modelos computacionais para *DNA shuffling* são úteis na complementação dos procedimentos experimentais e compreensão de questões referentes à frequência dos mutantes na diversidade de uma biblioteca. Alguns programas são citados a seguir, sendo que o detalhamento deverá ser encontrado em suas fontes originais.

O primeiro programa proposto (Moore et al., 1997) consistiu em um sistema matemático simples de recombinação para diferentes moléculas envolvidas em experimentos de *DNA shuffling*. No entanto, esse programa limita os tamanhos dos fragmentos a ser recombinados a valores menores que as distâncias entre as mutações. O modelo matemático desenvolvido por Sun (1999) pode ser aplicado nas situações onde se deseja o cálculo da quantidade do produto de *shuffling*, de acordo com um número determinado de mutações e/ou a construção do mapa físico usando *DNA shuffling*. O modelo utiliza, inicialmente, o método de Lander & Waterman (1988) usado com sucesso para modelar a distribuição de fragmentos clonados randomicamente ao longo do DNA genômico (mapeamento físico). No *DNA shuffling* o mapeamento é utilizado para visualizar a distribuição das regiões que serão recombinadas ao longo dos genes, provenientes de espécies diferentes. Quando os dados experimentais produzidos por Stemmer (1994a,b) e Zhao & Arnold (1997) foram utilizados para validação deste programa, observou-se a aproximação dos resultados obtidos com o modelo teórico (Sun, 1999).

O programa denominado eShuffle (Moore et al., 2001) modela a etapa de anelamento, através de informação da seqüência completa dos genes e dos cálculos de energia livre. Dados estatísticos, obtidos para os eventos de anelamento, são combinados com o algoritmo de recombinação inferindo localizações de mutações em seqüências recombinadas. O programa permite estimar a fração de seqüências recombinadas, a partir de um ou mais cruzamentos, e indica a probabilidade de determinados nucleotídeos ocorrerem em posições definidas e pertencerem ao sítio do evento de recombinação.

O eCodonOpt é um sistema que visa orientar posições de recombinações, gerando perfis da identidade e energia livre de um grupo de seqüências através da otimização do códon preferencial (códon de bases nucleotídicas preferidos por determinados organismos para codificação de biomoléculas) (Moore et al., 2002). O modelo computacional é constituído de fórmulas específicas para: (i) maximização do número médio de cruzamentos por seqüências recombinadas, (ii) minimização de tendências de recombinação de uma família de genes (por exemplo, formação de homoduplexes), e (iii) maximização da frequência relativa de recombinação em estruturas tridimensionais específicas (por exemplo, regiões de alças).

5. Aplicação da metodologia de *DNA shuffling* no melhoramento genético de plantas.

A aplicação da tecnologia do *DNA shuffling* permite o melhoramento de genes que poderão solucionar problemas específicos da agricultura, tais como o desenvolvimento de resistência a estresse biótico e/ou abiótico. Exemplos clássicos do emprego desta metodologia, na seleção de genes com características importantes, já foram relatados no melhoramento de marcadores moleculares, desenvolvimento de tolerância a herbicidas e resistência a insetos-praga.

5.1 Genes marcadores

O gene de β -glucuronidase (GUS) de *E coli* é um dos mais utilizados nas pesquisas envolvendo engenharia genética de plantas, funcionando freqüentemente, como marcador nos sistemas de transformação de plantas. A principal limitação deste sistema é a baixa atividade da enzima na presença dos fixadores formaldeído ou glutaraldeído. A fim de minimizar as limitações do sistema, Matsumura et al. (1999) submeteram o gene a três ciclos de recombinação usando *DNA shuffling* e seleção sob concentrações crescentes de glutaraldeído, resultando no isolamento de genes codificadores para GUS com tolerância melhorada em relação ao fixador. O melhor mutante apresentou aproximadamente 80% de atividade enzimática após tratamento com 0,2% de glutaraldeído. Posteriormente, estudos com o mesmo gene, isolado da bactéria termofílica (*Thermotima maritoga*), após dois ciclos de *shuffling* e seleção para atividade em 37°C identificaram uma proteína mutante melhorada em mais de 30 vezes em relação à proteína original, nesta temperatura (Lassner & McElroy, 2002). Resultados interessantes foram também relatados para o gene marcador que codifica para a proteína fluorescente GFP ("green fluorescence protein"), proveniente da "Água-viva" (Classe Scyphozoa). Variantes desta proteína apresentaram elevado aumento na atividade, após a aplicação da técnica de *DNA shuffling* (Crameri et al., 1996; Shivprasad et al., 1999).

5.2 Tolerância a herbicidas

Os herbicidas são amplamente utilizados na agricultura mundial com a finalidade de controlar o desenvolvimento de ervas daninhas, além de reduzir custos de produção através da minimização do uso de mão-de-obra. Os mais utilizados pertencem à classe das triazinas e glifosatos, sendo a descoberta de enzimas, que os degradam, de extrema importância no desenvolvimento de plantas resistentes aos herbicidas, evitando a detecção de resíduos que poderiam gerar problemas na comercialização de grãos. Estudos empregando metodologia de *DNA shuffling* foram realizados para enzimas que degradam herbicidas, a fim de produzir novas moléculas com atividades melhoradas. Com este objetivo, construiu-se uma biblioteca combinatória de mutantes para os genes de triazinas hidrolases (*Atz A* e *TriA*, efetivas contra cloro e amina, respectivamente) (Raillard et al., 2001) e o produto, contendo 1.600 recombinantes, foi funcionalmente selecionado contra 15 grupos químicos de triazinas, resultando no incremento de até 150 vezes na atividade enzimática e na detecção de moléculas com capacidade de degradar até 7 grupos químicos do herbicida. Com o mesmo intuito, He et al. (2001) recombinaram genes de duas fontes diferentes, que codificam para a enzima EPSPS (*5-enolpiruvilshikimate-3-fostato sintase*, responsável pela degradação do glifosato), resultando no aumento da degradação do herbicida e no melhoramento dos parâmetros cinéticos das novas ESPPS desenvolvidas. Estratégias como estas podem prover novas soluções no melhoramento protéico funcional, anteriormente obtido apenas por mutagênese randômica e/ou mutagênese baseada em predição.

5.3 Resistência a insetos

As proteínas de defesa dividem-se em três classes, conforme função desempenhada. A primeira classe agrupa produtos capazes de mudar diretamente as propriedades extracelulares da matriz, aumentando a defesa por reforçar, reparar ou alterar a parede celular. O grupo inclui proteínas estruturais constituídas de glicoproteínas ricas em hidroxiprolina/glicina e enzimas envolvidas na construção/modificação de polímeros da parede celular (lignina, suberina, compostos fenólicos e calose) (Gomes & Xavier-Filho, 1994). O segundo grupo contém proteínas com atividade antimicrobiana direta, exemplificando-se hidrolases pertencentes a quitinase e α -1,3-glucanase, catalisadores da síntese de antimicrobianos ou contra herbívoros (inibidores de proteinases e α -amilases), proteínas tóxicas (lectinas, tioninas ou inativadores de ribossomos e proteínas Cry de *B. thuringiensis*), além de enzimas envolvidas na síntese de metabólitos secundários como taninos, quinonas e fitoalexinas com atividade antimicrobiana (Bowles, 1990; Franco et al., 2002). A terceira classe envolve proteínas com resposta de defesa relacionada com patogenicidade (PR) (Gomes & Xavier-Filho, 1994).

As pesquisas vêm se acelerando nos últimos anos visando identificar novas substâncias produzidas por plantas e microrganismos ou melhorar geneticamente as já existentes, capazes de participar direta ou indiretamente nas estratégias de defesa contra insetos e patógenos. Estas substâncias pertencem ao grupo de fenóis, aminoácidos não protéicos, alcalóides, lectinas, inibidores de proteinases e de α -amilases (Gatehouse et al., 1992), proteínas Cry de *B. thuringiensis* (Marzari et al., 1997), toxinas “killer” produzidas por leveduras (Weiler & Schmitt, 2003), entre outras.

A produção de plantas transgênicas expressando proteínas Cry e inibidores enzimáticos, que atuam no controle de pragas de importância econômica na agricultura, representa um dos principais enfoques da biotecnologia atual. Nesse contexto, pesquisas na Embrapa Recursos Genéticos e Biotecnologia vêm sendo desenvolvidas utilizando bibliotecas de mutantes de proteínas Cry, de inibidores de α -amilases e de inibidores de proteinases, com objetivo de prospectar genes codificadores de moléculas melhoradas, efetivas no controle de insetos-praga para importantes culturas.

5.3.1 Biblioteca combinatória de proteínas Cry

As proteínas Cry apresentam atividades específicas e restritas a determinados insetos-praga, com a vantagem de apresentar inocuidade aos vertebrados, principalmente aos humanos e ao meio ambiente (Crickmore et al., 1998), fazendo deste agente um componente-chave no controle de insetos-praga. O uso de técnicas moleculares na obtenção de toxinas melhoradas (*DNA shuffling*) e construção de biblioteca combinatória apresentada em fagos (*Phage display*), permitem selecionar toxinas com atividade dirigida ao inseto alvo, contribuindo para a segurança quanto ao uso de plantas geneticamente modificadas. Nossa equipe, na unidade de pesquisa da Embrapa Recursos Genéticos e Biotecnologia, vem desenvolvendo pesquisas visando à obtenção de bibliotecas combinatórias de genes *cry* para seleção de proteínas ativas contra pragas, em particular contra o bicudo do algodoeiro (*Anthonomus grandis*). O bicudo do algodoeiro é considerado, atualmente, uma das pragas de maior relevância da agricultura mundial, devido aos danos causados e dificuldades de controle. O combate desse inseto pelos métodos convencionais com a aplicação de pesticidas, além de oneroso e prejudicial ao meio ambiente é ineficiente, devido à fase larval do patógeno ser endofítica (desenvolvimento no interior do botão floral ou das maçãs). Portanto, o uso de bibliotecas de genes recombinados na seleção de toxinas com atividade específica para o bicudo do algodoeiro poderá contribuir de forma marcante no desenvolvimento de plantas geneticamente modificadas resistentes a insetos.

5.3.2 Biblioteca cominatória de inibidores de α -amilase

Os inibidores de α -amilases encontrados em sementes de plantas como trigo, feijão, centeio, milho, entre outros, são conhecidos por conferir resistência aos insetos. Em particular, o inibidor de α -amilase (α AI-1), presente em feijão comum (*Phaseolus vulgaris*) e a isoforma α AI-2, presente no feijão selvagem apresentam diferentes especificidades: enquanto o α AI-1 inibe a α -amilase pancreática de porco e as α -amilases dos bruquídeos *Calossobruchus maculatus* e *Calossobruchus chinensis*, o inibidor α AI-2 inibe apenas a α -amilase de *Zabrotes subfasciatus* (Grossi-de-Sá et al., 1997). Com o objetivo de entender a interação molecular entre α -amilases e inibidores de α -amilases de plantas, identificar resíduos de aminoácidos responsáveis pelas diferenças de especificidade entre as isoformas (Grossi-de-Sá et al., 1997; Franco et al., 2002; Silva et al., 2000; Silva et al., 2004) e desenvolver novos inibidores com atividade contra pragas de interesse econômico, aplicamos a tecnologia do *DNA shuffling* seguida de *Phage display* para a geração de uma biblioteca de genes mutantes de inibidores de α -amilases (α AI-1 e α AI-2). A biblioteca combinatória contendo 10^6 mutantes está sendo avaliada para a seleção de inibidores novos e/ou melhorados. O estudo estrutural e funcional de mutantes selecionados, em complexo com as α -amilases dos insetos, possibilitarão a identificação de aminoácidos essenciais para o desenho de outros inibidores específicos contra pragas danosas à agricultura, com conseqüente aplicação na produção de plantas transgênicas resistentes a insetos-praga.

Referências Bibliográficas

BOWLES, D. J. Defense-related proteins in higher plants. **Annual Review of Biochemistry**, Palo Alto, v. 59, p. 873-907, 1990.

CALDWELL, R. C.; JOYCE, G. F. Randomization of genes by PCR mutagenesis. **PCR Method Applications**, v. 2, p. 28-33, 1992.

CHEN K.; ARNOLD, F. H. Tuning the activity of an enzyme form unusual environments: sequential random mutagenesis of subtilisin E for catalysis in dimethylformamide. **Proceedings of the National Academy of Sciences**, Washington, v. 90, p. 5618-5622, 1993.

COCO, W. M.; ENCELL, L. P.; LEVINSON, W. E.; CRIST, M. J.; LOOMIS, A. K.; LICATO, L. L.; ARENSDORF, J. J.; SICA, N.; PIENKOS, P. T.; MONTICELLO, D.J. Growth factor engineering by degenerate homoduplex gene family recombination. **Nature Biotechnology**, New York, v. 20, p. 1246-1250, 2002.

CRAMERI, A.; WHITEHORN, E. A.; TATE, E. Improved green fluorescent protein by molecular evolution using DNA shuffling. **Nature Biotechnology**, New York, v. 14, p. 315-319, 1996.

CRAMERI, A.; RAILLARD, S. A.; BERMUDEZ, E.; STEMMER, W. P.; DNA shuffling of a family of genes from diverse species accelerates directed evolution. **Nature**, London, v. 391, p. 288-291, 1998.

CRICKMORE, N.; ZEIGLER, D. R.; FEITELSON, J.; SCHNEPF, E.; VANRIE, J.; LERECLUS, D.; BAUM, J.; DEAN, D. H. Revision of the nomenclature for the *Bacillus thuringiensis* pesticidal crystal proteins. **Microbiology and Molecular Biology Reviews**, Washington, v. 62, p. 807-813, 1998.

FRANCO, O. L.; RIDGEN, D. J.; MELO, F.; GROSSI-DE-SÁ, M. F. Plant α -amylase inhibitors and their interaction with insect α -amylases. **European Journal of Biochemistry**, New York, v. 269, p. 1-17, 2002.

GATEHOUSE, A. M. R.; HILDER, V. A.; GATEHOUSE, J. A. Control of insect pests by plant genetic engineering. **Proceedings of the Royal Society of Edinburgh**, Edinburgh, v. 99B, n. (3/4), p. 51-60, 1992.

GOMES, V. M.; XAVIER-FILHO, J. Biochemical defenses of plants. **Arquivos de Biologia e Tecnologia**, Curitiba, v. 37, n. 2, p. 371-383, 1994.

GROSSI-DE-SÁ, M. F.; MIRKOV, T. E.; ISHIMOTO, M.; COLUCCI, G.; BATEMEM, K.; CHISPEELS, M. J. Molecular characterization of a bean alpha amylase of the Mexican bean weevil *Zabrotes subfasciatus*. **Planta**, New York, v. 203, p. 295-303, 1997.

HE, M.; YANG, Z.; NIE, Y. A new type of class 1 bacterial 5-enolpyruvylshikimate-3-phosphate synthase mutants with enhanced tolerance to glyphosate. **Biochemica et Biophysica Acta**, Amsterdam, v. 1568, p. 1-6, 2001.

KIKUCHI, M.; OHNISHI, K.; HARAYAMA, S. Novel family shuffling methods for the in vitro evolution enzymes. **Gene**, Amsterdam, v. 236, p. 159-167, 1999.

KIKUCHI-NODATE, M.; KAGAMI, O. Molecular Breeding of useful enzymes by novel DNA shuffling methods. Available in <http://www.salmon.mbio.co.jp/mbi/english/kamaishi/program/DNAshuffli/DNAshuffli.html>>. Access on April, 2002.

KIKUCHI, M.; OHNISHI, K.; HARAYAMA, S. An effective family shuffling method using single-stranded DNA. **Gene**, Amsterdam, v. 243, p. 133-137, 2000.

LANDER, E. S.; WATERMAN, M. S. Genomic mapping by fingerprinting random clones: a mathematical analysis. **Genomics**, San Diego, v. 2, p. 231-239, 1988.

LASSNER, M. The evolution of DNA shuffling. In: INTERNATIONAL CONGRESS OF MOLECULAR BIOLOGY, 7., 2003. Barcelona, Spain. [**Proceedings**. Barcelona]: Molecular Biology Society, 2003. p. 249. S18-1.

LASSNER, M.; McELROY, D. Directed molecular evolution: Bridging the gap between genomics leads and commercial products. **OMICS A Journal of Integrative Biology**, Lanchmont, v. 6, p. 153-162, 2002.

LUTZ, S.; OSTERMEIER, M.; MOORE, G. L.; MARANAS, C. D.; BENKOVIC, S. J. Creating multiple-crossover DNA libraries independent of sequence identity. **Proceedings of the National Academy of Sciences**, Washington, v. 98, p. 11248-11253, 2001.

MARZARI, R.; EDOMI, P.; BHATNAGAR, R.K.; AHMAD, S.; SELVAPANDIYAN, A.; BRADBURY, A. Phage display of *Bacillus thuringiensis* CryIA(a) insecticidal toxin. **FEBS Letters**, Amsterdam, v.411, p.27-31, 1997.

MATSUMURA, I.; WALLINGFORD, J.; SURANA, N. Directed evolution of the surface chemistry of the reporter enzyme b-glucuronidase. **Nature Biotechnology**, New York, v. 17, p. 696-701, 1999.

MOORE, C. M.; JIM, H. M.; KUCHNER, O.; ARNOLD, F. H. Strategies for the in vitro evolution of protein function: enzyme evolution by random recombination of improved sequences. **Journal of Molecular Biology**, London, v. 272, p. 336-347, 1997.

MOORE, G. L.; MARANAS, C. D.; LUTZ, S.; BENKOVIC, J. Predicting crossover generation in DNA shuffling. **Proceedings of the National Academy of Sciences**, Washington, v. 98, p. 3226-3231, 2001.

MOORE, G. L.; MARANAS, C. D. eCodonOpt: a systematic computational framework for optimizing codon usage in directed evolution experiments. **Nucleic Acids Research**, Oxford, v. 30, p. 2407-2416, 2002.

NEDWIN, G. E. Directed evolution of a fungal peroxidase. In: SAYLER, G. (Ed.). **Biotechnology in the sustainable environment**. New York: Plenum Press, 1997. p. 13-32.

NESS, J. E.; KIM, S.; GOTTMAN, A.; PAK, R.; KREBBER, A.; BORCHERT, T. V.; GOVINDARAJAN, S.; MUNDORFF, E. C.; MINSHULL, J. Synthetic shuffling expands functional protein diversity by allowing amino acids to recombine independently. **Nature Biotechnology**, New York, v. 20, p. 1251-1255, 2002.

OSTERMEIER, M. Synthetic gene libraries: in search of the optimal diversity. **Trends in Biotechnology**, Barking, v. 21, p. 244-247, 2003.

RAILLARD, S.; KREBBER, A.; CHEN, Y. Novel enzyme activities and functional plasticity revealed by recombining highly homologous enzymes. **Chemical Biology**, Cambridge, v. 8, p. 891-898, 2001.

REIDHAAR-OLSON J.; BREYER, R. M.; HU, J. C.; KNIGHT K. L.; LIM, W. A.; MOSSING, M. C.; PARSELL, D. A.; SHOEMAKER, K. R.; SAUJER, R. T. Random mutagenesis of protein sequences using oligonucleotide cassettes. **Methods in Enzymology**, San Diego, v. 208, p. 564-588, 1994.

SHIVPRASAD, S.; POGUE, G. P.; LEWANDOWSKI, D. J. Heterologous sequences greatly affect for eign gene expression in tobacco mosaic virus-based vectors. **Virology**, San Diego, v. 255, p. 312-323, 1999.

SILVA, M. C. M da.; GROSSI-DE-SÁ, M. F.; CHRISPEELS, M. J.; TOGAWA, R. C.; NESHICH, G. Analysis of structural and physico-chemical parameters involved in the specificity of binding between alpha-amylases and their inhibitors. **Protein Engineering**, Oxford, v.13, p. 167-177, 2000.

SILVA, M. C. M. da; MELLO, L. V.; COUTINHO, M. V.; RIDGEN, D. J.; NESHICH, G.; CHRISPEELS, M. J.; GROSSI-DE-SÁ, MF. Bean a-amylase inhibitor mutants to investigate specificity of binding to a-amylases. **Pesquisa Agropecuária Brasileira**, Brasília, v. 39, p. 201-208, 2004.

STEMMER, W. P. C. Rapid evolution of a protein in vitro by DNA shuffling. **Nature**, London, v. 370, p. 389-391, 1994a.

STEMMER, W. P. C. DNA shuffling by random fragmentation and reassembly: *in vitro* recombination for molecular evolution. **Proceedings of the National Academy of Sciences**, Washington, v. 91, p. 10747-10751, 1994b.

STOOP, A. A.; JESPERS, L.; LASTERS, I.; ELDERING, E.; PANNEKOEK, H. High-density mutagenesis by combined DNA shuffling and Phage Display to assign essential amino acid residues in protein-protein interactions: application to study structure-function of plasminogen activation inhibitor 1 (PAI-I). **Journal of Molecular Biology**, London, v. 301, p. 1135-1147, 2000.

SUN, F. Modeling DNA shuffling. **Journal of Computational Biology**, Rockville, v. 6, p. 77-90, 1999.

VIEIRA, J.; MESSING, J. Production of single-stranded plasmid DNA. **Methods in Enzymology**, San Diego, v. 153, p. 3-11, 1987.

WEILER, F.; SCHMITT, M. J. Zygotin, a secreted antifungal toxin of the yeast *Zygosaccharomyces bailii*, and its effect on sensitive fungal cells. **FEMS Yeast Research**, Amsterdam, v. 3, n. 1, p. 69-76, 2003.

YANO, T.; OUE, S.; KAGAMIYAMA, H. Directed evolution of an aspartate aminotransferase with new substrate specificities. **Proceedings of the National Academy of Sciences**, Washington, v. 95, p. 551-5515, 1998.

YOU, L.; ARNOLD, F. H. Directed evolution of subtilisin E in *Bacillus subtilis* to enhance total activity in aqueous dimethylformamide. **Protein Engineering**, Oxford, v. 9, p. 77-83, 1994.

ZHANG, J. H.; DAWES, G.; STEMMER, W. P. Directed evolution of a fucosidase from a galactosidase by DNA shuffling and screening. **Proceedings of the National Academy of Sciences**, Washington, v. 94, p. 4504-4509, 1997.

ZHAO, H.; ARNOLD, F. H. Optimization of DNA shuffling for high fidelity recombination. **Nucleic Acids Research**, Oxford, v. 25, p. 1307-1308, 1997.