

SAGE Suite

Uma suite para análise e geração de dados *in-silico* de SAGE

O que é SAGE?

SAGE (Serial Analysis of Gene Expression) é uma técnica para quantificar em larga escala a expressão de genes de uma dada população de RNAs mensageiros. Esta técnica gera etiquetas (*tags*) de cDNA que são concatenadas e seqüenciadas para, em seguida, serem analisadas e anotadas.

A Técnica

1. Isolamento do RNAm

- O RNA mensageiro é isolado do resto do material gênico
- O RNAm é completado em cDNA

A Técnica

2. Geração das Tags

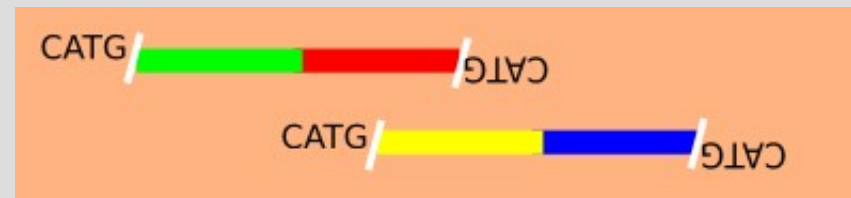
- Utilizando um enzima de restrição os sitios CATG são cortados
- Uma segunda enzima de restrição corta as seqüências 10 pb (18 pb no LongSAGE) a frente do sítio CATG
- As Tags de são isoladas do resto do material gênico



A Técnica

3. Formação de Ditags

- As Tags são separadas em dois conjuntos.
- Com o uso de dois adaptadores as Tags de cada conjunto são unidas, duas a duas.
- Particularidade: Em Long SAGE as Tags não são cortadas homogeneamente. O final possui uma diferença de duas bases, adicionando assim um viés à união.



A Técnica

4. União em Concatâmeros e seqüenciamento

- As Ditags são unidas em grandes seqüências, chamadas concatâmeros e, finalmente, seqüenciadas.



Problemas

- ♦ Imprecisão da enzima de corte de Long SAGE (tags de 17 ou 18 pb)
- ♦ Erros de sequenciamento
- ♦ Contaminação da amostra
- ♦ Degeneração das seqüências

Geração de Dados

Motivação

- A confiabilidade dos programas existentes é dúbia, logo foi criado um sistema próprio
- Existe a necessidade de se testar o programa com dados reais
- Não se conhece o resultado “ideal” de dados reais

Geração de Dados

1. Geração de contagens

- Uma lista de Tags aleatórias é gerada
- Para cada Tag uma contagem é sorteada
- Diversas distribuições podem ser utilizadas para gerar expressões gênicas variadas

Geração de Dados

2. Corte das bases finais

- ♦ Simula a instabilidade da enzima de Long SAGE
- ♦ Dada uma probabilidade de corte sorteia aleatoriamente para cada Tag, gerando assim Tags com tamanhos variados

Geração de Dados

3. União em Dtags

- Uni as Tags duas a duas aleatoriamente.
- Para o caso de Long SAGE, extremidade coesiva, o sorteio é feito de maneira diferente, levando em conta o viés gerado.

Geração de Dados

4. Formação de Concatameros

- Agrupa as Ditags em concatameros
- Gera dentro de um tamanho mínimo e máximo estipulado
- É possível introduzir erro ao sinal de pontuação para avaliar a capacidade de recuperação do analisador

Interface

Versão Gráfica local

- Construída em perl/Tk
- Facilita e agiliza o trabalho do usuário
- Permite maior interação com o usuário

Versão Web

- Independe da instalação de qualquer programa
- Pode ser feita de qualquer local com facilidade

Trabalhos Futuros

- Finalização da interface gráfica para o SAGE Analysis
- Término do desenvolvimento da interface web e alocação em servidor
- Análise comparativa com programas já existentes