

名前

miranda - ゲノム配列から miRNA のターゲットとなる可能性のある部位を探し出します。

書式

基本的な使い方:

```
miranda file1 file2 [ options ... ]
```

応用:

```
miranda file1 file2 [-sc score] [-en energy]
[-scale scale] [-loose] [-go X] [-ge Y] [-out fileout]
[-quiet] [-trim T] [-noenergy] [-shuffle] [-s nshuff]
[-w W] [-uniform] [-z zscore]
```

説明

miRanda は、ゲノム配列から microRNA のターゲットとなる可能性のある部位を検出するためのアルゴリズムです。**miRanda** は、**file1** から (microRNA のような) RNA 配列を、**file2** からはゲノムの DNA/RNA 配列を読み込みます。両方のファイルともに、FASTA フォーマットでなければなりません。

FASTA フォーマットの例:

```
>embl|AJ550546|DME550546 Drosophila melanogaster microRNA miR-bantam
GTGAGAUCAUUUUGAAAGCUG
```

file1 より読み出された一つもしくは複数の miRNA 配列について、**file2** の全ての配列に対してスキャンが行われます。ターゲットの可能性のある部位は、2 ステップのストラテジーによって同定されます。最初に、miRNA の質問配列と参照配列との間で、ダイナミック・プログラミング・ローカル・アライメントが実行されます。このアライメントによるスコアは配列の相補性に基づくものであり、配列の同一性によるものではありません。言い換えると、A:A や G:G の代わりに、A:U や G:C のような塩基対のマッチを見つけ出します。G:U wobble base pair (ゆらぎ塩基対) も許容されますが、通常、最適マッチの場合よりも低いスコアが加算されます。

アライメントの例：

```
Query:    3' GTCGAAAGTTTTACTAGAGTG 5' (eg. miRNA)
          :||:||||| |||||
Ref:      5' TAGTTTTTCACAATGATCTCGG 3' (eg. 3'UTR)
```

アルゴリズムの第 2 段階は、第 1 段階で高いスコアを得たアライメント (-sc によって定義された閾値以上のスコアを持つもの) に対して行われ、RNA 二重鎖の熱力学的な安定性を基準にこれらのアライメントを評価します。この第 2 段階の評価法では、Ivo Hofacker によって開発された **ViennaRNA** パッケージの一部である **RNALib** ライブラリーの折りたたみルーチンを利用しています。まずこの段階では、質問配列、リンカー、参照配列(逆転される)から構成されている、強制的な仮想の一本鎖 RNA を生成します。次に、**RNALib** を用いて、この構造が折りたたまれます。さらにオプションとして、シャッフルされた参照配列 (-shuffled オプションを参照) を用いた多数のアライメントの実行によって、それぞれのターゲット配列についての初歩的な統計値を算出します。

これらのデータから分布が構築され、そしてこの分布から得られた統計学的なパラメータは、検出されたターゲット配列の Z-スコアを算出するのに使われます。

オプション

--help -h

ヘルプ、利用方法に関する情報、コマンドラインのオプションを表示する

--version -v --license

バージョン情報、ライセンス情報を表示する

-sc score

スコアの閾値を **score** に設定。この値以上のスコアのアライメントのみについて、その後の解析が行われる。

-en energy

エネルギーの閾値を **energy** に設定。この値以下のアライメントのみについて、その後の解析が行われる。この値は、負を取る。

-scale scale

スケーリング・パラメータを **scale** に設定。このスケーリングは、決定的な意味を持つ microRNA の 5'末端の 10bp 領域のマッチ/ミスマッチに対して適用される。多くの既知の miRNA:ターゲット二本鎖では、この領域において相補性の高いことが知られている。このパラメータは、より効果的にこの種のアライメントを検出するための、コントラスト機能と考えることができる。

-loose

厳密なアライメントのヒューリスティック（発見的手法）を除外する。通常モードでは、ヒューリスティック的な手法が、二重鎖の予想で使われている：既知の miRNA:ターゲット二重鎖の知見に基づいて、異なるアライメント領域での相補的塩基対が数えられ、このテストに通らなかったアライメントは表示されない。このオプションを off にした場合、安全性が低く、擬陽性の検出結果が増える。

-go X

アライメントのギャップ・オープニング・ペナルティの値を **X** に設定。この値は負数である。

-ge Y

ギャップ延長ペナルティの値を **Y** に設定。この値は負数である。

-out fileout

結果を **fileout** という名前の出力ファイルに書き出す。

-quiet

Quiet モード。出力を最小にする。

-trim T

参照配列から **T** ヌクレオチドを切り取る。参照配列の 3'UTR が noisy である事が事前に予想される場合に便利。

-shuffle

それぞれの microRNA と参照配列との間の解析について、microRNA とシャッフルされた参照配列との間で、**nshuff** で設定された回数のアライメントを実行する。このアライメントは、与えられたスコアの信頼性を見積もる時に、シャッフルされたアライメントのスコア分布を生成する目的で行われる。それぞれの検出されたターゲットの **Z-Score** がこの解析によって生成される。**-z** オプションによって、このスコアによる結果のフィルタリングをすることができる。ただし、これによって劇的にアライメントのプロセス数が増加し、アルゴリズム的に非常に計算が遅くなることに注意せよ。

-s nshuff

ランダム・シャッフル解析の全体数を **nshuff** に設定する。この値を大きくすると、より信頼性の高い Z-score が得られる。

-w W

ウィンドウ・ベースのシャッフルをおこなう際、**W** パラメーターを変更することで、ウィンドウ・サイズが変更になる。ウィンドウ・ベースのシャッフルは、配列のヌクレオチド組成を

保ちつつ、配列上を動く小さなウィンドウ内で行われる。

-uniform

ウィンドウ・シャッフルを行わず、参照配列上を均一にシャッフルする。

-z zscore

全てのターゲットに対し、配列のシャッフルの統計学に基づく **Z-Score** の閾値を設定する。
Z-Score が **zscore** 以上の値をとるもののみが検出されて、アルゴリズムに基づき表示される。

参考文献

もし、このプログラムを研究に用いる場合は、下記の引用文献も含めること：

A.J. Enright, B. John, U. Gaul, T. Tuschl, C. Sander, D.S. Marks; (2003) MicroRNA targets in Drosophila; *Genome Biology* 5(1):R1.

RNAlib Citations:

I.L. Hofacker, W. Fontana, P.F. Stadler, S. Bonhoeffer, M. Tacker, P. Schuster (1994) Fast Folding and Comparison of RNA Secondary Structures. *Monatshefte f. Chemie* 125: 167-188

M. Zuker, P. Stiegler (1981) Optimal computer folding of large RNA sequences using thermodynamic and auxiliary information, *Nucl Acid Res* 9: 133-148

J.S. McCaskill (1990) The equilibrium partition function and base pair binding probabilities for RNA secondary

バグ

ご意見、バグレポート等は、miranda@cbio.mskcc.org へ。

Anton Enright

1.0

miranda(l)

This document was translated by Mitsuhiro Tagaya(magaya@tetralog.mydns.jp) in 29 July, 2005.