

The NextGENe Barcode Sorting Tool

データファイルにバーコード配列が含まれる場合は、NextGENe Barcode Sorting ツールを使って解析前にバーコード配列が付加されたリードデータを別々のファイルに分類してください。Barcode Sorting ツールを使うと、バーコード付加されたサンプルファイルを配列タグに基づいて別々のファイルに分類できます。このツールは、シーケンスリードにバーコード配列が付加されているデータファイルやバーコードがリード名もしくは別のファイルに含まれているデータファイルに使用できます。リードからこれらのタグをトリミングして、それらのリードをタグに基づいて区別するために、以下の 2 つのオプションが使用できます；

- 全バーコードの詳細（バーコード配列タグとそのサンプル ID）がわかっている場合、タブ区切りテキストファイルでバーコード/プライマーファイルを作成して、NextGENe Barcode Sorting ツールにサンプル ID やフォワードバーコード/プライマータグ、リバースバーコード/プライマータグの情報を与えることができます。
- バーコードの詳細が不明な場合、NextGENe Barcode Sorting ツールを使って自動的にバーコード配列タグやタグ総数を検出してから、各タグのフォルダを作成できます。

Barcode/Primer File

Excel などのソフトを使ってバーコード/プライマーファイルを作成できます。作成したらタブ区切りテキストファイルで保存してください。ファイルの各行にはサンプル ID とそのバーコードタグのエントリーを含めてください。図 8-1 はあるサンプルバーコード/プライマーファイルです。各サンプルに 2 つのタグが付加されています。ファイルの各行には、サンプル ID (Sample_ID)、フォワードバーコードタグ (Forward Tag)、リバースバーコードタグ (Reverse Tag) が含まれます。

図 8-1 : 2 タグのバーコード/プライマーファイル例

Sample ID	Forward Tag	Reverse Tag
AYR_1A	GTGAGGCTTGTCTCAAAGATTAAGCC	GTGAGGCCTGCTGCCTTCCTTGA
AYR_1B	TACGCGCTTGTCTCAAAGATTAAGCC	TACGCGCCTGCTGCCTTCCTTGA
AYR_1C	GTCACGCTTGTCTCAAAGATTAAGCC	GTCACGCCTGCTGCCTTCCTTGA

✓ リバースタグが使われていない場合は、リバースタグ列をブランクにしてください。

図 8-2 は各サンプルに複数のタグが含まれるバーコード/プライマーファイルのサンプルです。

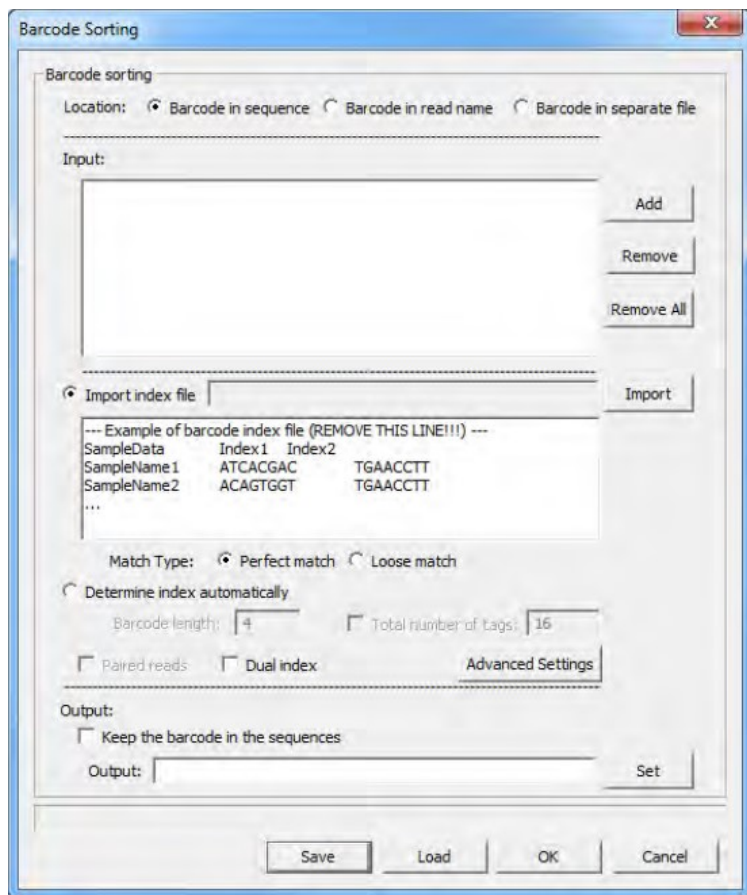
図 8-2 : 複数タグのバーコード/プライマーファイル例

Sample ID	Tag 1	Tag 2	Tag 3	Tag 4
Samp12345	CCGTGAACGT	CCGTTAACGT	CCGTGACCCT	CCGTAACTG
Samp23456	CCGTGACCGT	CCGTACCCT	CCGTGATGAC	CCGTGAGTAC

To parse barcoded sample files

1. NextGENe メインメニュー Tools メニュー>Barcode Sorting をクリックしてください。
Barcode Sorting ダイアログボックスが開きます。

図 8-3 : Barcode Sorting ダイアログボックス



2. サンプルファイルの Barcode Sorting オプション>Location (Barcode in sequence、Barcode in read name、Barcode in separate file) を指定してください。

Barcode in read name ダイアログボックスのオプションは、選択した Barcode Sorting オプションによって変わります。

3. Add をクリックしてサンプルファイルを選択してください。

サンプルファイルは Sample List ペインにファイル名順にリストアップされます。ファイル名には各サンプルファイルのフルディレクトリパスが含まれます。

4. 以下のオプションのどちらかを選択してください；

設定	概要
<p>Import index file</p>	<p>既知のバーコード情報でバーコード/プライマーファイルを作成している場合、このオプションを選択してください。Import をクリックしてインポートするバーコード/プライマーファイルを選択してから以下のどちらかを選択してください；</p> <ul style="list-style-type: none"> ・ Perfect Match・・・このオプションを選択する場合、リードのタグはバーコード/プライマーファイル内で定義されているタグと完全に一致していなければいけません。完全一致でない場合、リードはそのタグに割り当てられません。 ・ Loose Match・・・このオプションを選択する場合、リードのタグは、前半・後半・中央セグメントに分けられ、これらのセグメントのうち一つがバーコード/プライマーファイル内のタグと一致すればリードはそのタグに割り当てられます。 ✓ Loose Match 法は、タグ配列中のシーケンスエラーの可能性が大きい、比較的長いタグ配列に特に有効です。
<p>Determine index automatically</p>	<p>バーコード情報が不明で、自動的にバーコード情報を検出したい場合はこのオプションを選択してから、以下を行ってください；</p> <ul style="list-style-type: none"> ・ バーコード長の設定（Barcode in sequence を選択した場合のみ使用できます。） ・ タグの総数がわかっている場合は、Total Number of Tags を選択してからその値を入力してください。

5. 選択した Barcode Sorting オプションに応じて、以下を行ってください；

選択した Barcode Sorting	オプション
Barcode in sequence	<ol style="list-style-type: none">1. フォワードリードとリバースリードがそれぞれバーコードをもつ場合、Dual index を選択してください。2. Step 6 に進んでください。
Barcode in read name	Step 8 に進んでください。
Barcode in separate file	<ol style="list-style-type: none">1. 以下の一つもしくは両方を行ってください；<ul style="list-style-type: none">・ Paired Reads : ペアリードデータをロードしている場合、Paired Reads を選択してください。・ Dual index : フォワードリードとリバースリードがそれぞれバーコードをもつ場合、Dual index を選択してください。2. Step 8 に進んでください。

6. Output ペインで、以下を行ってください；

- Barcode in sequence を選択しており、出力ファイル内のリードにバーコード配列を含める場合は、Keep the barcode in the sequences を選択してください。
- 出力ファイルの保存場所を指定してください。

7. OK をクリックしてください。





処理が完了したことを示すメッセージが開きます。Determine Automatically を選択して Total tag count を指定しなかった場合、タグ配列によるソーティングが完了する際、以下の 2 つの相互的な除外クライテリアが決定に使用されます。

- サンプルタグに含まれるリードカウントが、前のタグのカウントの 10%より小さいとき、そのタグは使われずに Barcode Sorting は完了します。
- サンプルリードの 95%がバーコードによって分類された後、1 つの追加タグがソーティングに使用され、ソーティングが完了します。

分類により生成された各データファイル名には、以下の情報が付加されています；

- タグ情報 (Determine Automatically が選択された場合)
- サンプル ID (バーコード/プライマーファイルが使用された場合)

図 8-5 : NextGENe Barcode Sorting ツールによって生成された各データファイル

 SRR018422_converted_CGAG.fasta	1/27/2010 10:51 AM	FASTA File	662 KB
 SRR018422_converted_CTCG.fasta	1/27/2010 10:51 AM	FASTA File	334,720 KB
 SRR018422_converted_OtherTags.fasta	1/27/2010 10:51 AM	FASTA File	8,509 KB
 SRR018422_converted_TCGA.fasta	1/27/2010 10:51 AM	FASTA File	1,166 KB

お問い合わせ先

電話・Eメールでのお問い合わせ

- バイオアップロード合同会社
- TEL : 0284-22-4213
- E-mail : info@bio-upload.com
- 対応時間帯 : 平日 9 : 00 ~ 17 : 30