Ultrafast and memory-efficient alignment of short DNA sequences to the human genome

Ben Langmead, Cole Trapnell,
Mihai Pop and Steven L Salzberg

Akshay Sanjeev

May 4, 2025

We are in 2009

Outline of Bowtie

- Burrows-Wheeler Transform(Indexing)
- Exact and inexact alignment
- Excessive backtracking

Burrows- Wheeler Transform: Forward Transform

Let $T = \texttt{BANANA}$. BWT(T) will be:

$$\begin{bmatrix}
\texttt{\$} & \texttt{B} & \texttt{A} & \texttt{N} & \texttt{A} & \texttt{N} & \texttt{A} \\
\texttt{A} & \texttt{\$} & \texttt{B} & \texttt{A} & \texttt{N} & \texttt{A} & \texttt{N} \\
\texttt{N} & \texttt{A} & \texttt{\$} & \texttt{B} & \texttt{A} & \texttt{N} & \texttt{A} \\
\texttt{A} & \texttt{N} & \texttt{A} & \texttt{\$} & \texttt{B} & \texttt{A} & \texttt{N} \\
\texttt{N} & \texttt{A} & \texttt{N} & \texttt{A} & \texttt{\$} & \texttt{B} & \texttt{A} \\
\texttt{A} & \texttt{N} & \texttt{A} & \texttt{N} & \texttt{A} & \texttt{\$} & \texttt{B} \\
\texttt{B} & \texttt{A} & \texttt{N} & \texttt{A} & \texttt{N} & \texttt{A} & \texttt{\$}
\end{bmatrix}$$

Burrows- Wheeler Transform: Forward Transform

Let $T = \texttt{BANANA}$. BWT(T) will be:

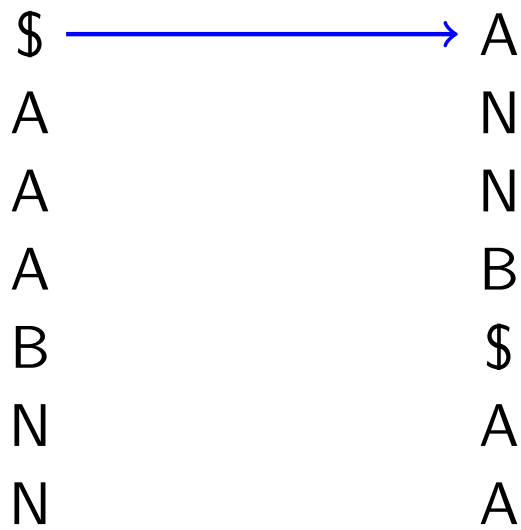
$$\begin{bmatrix} \$ & B & A & N & A & N & A \\ A & \$ & B & A & N & A & N \\ N & A & \$ & B & A & N & A \\ A & N & A & \$ & B & A & N \\ N & A & N & A & \$ & B & A \\ A & N & A & N & A & \$ & B \\ B & A & N & A & N & A & \$ \end{bmatrix} \rightarrow \begin{bmatrix} \$ & B & A & N & A & N & A \\ A & \$ & B & A & N & A & N \\ A & N & A & \$ & B & A & N \\ A & N & A & N & A & \$ & B \\ B & A & N & A & N & A & \$ \\ N & A & \$ & B & A & N & A \\ N & A & N & A & \$ & B & A \end{bmatrix}$$

$$BWT(T) \rightarrow \texttt{ANNB\$AA}$$

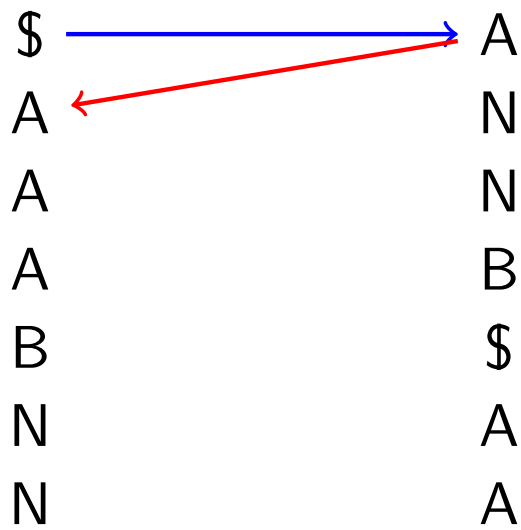
BWT: Reverse Transform(UNPERMUTE)

$	A
A	N
A	N
A	B
B	$
N	A
N	A

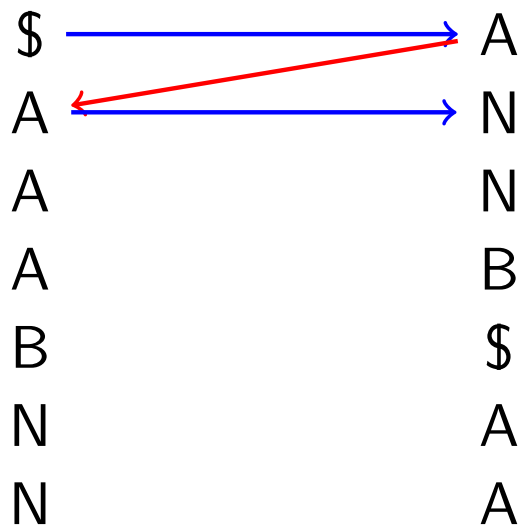
BWT: Reverse Transform(UNPERMUTE)



BWT: Reverse Transform(UNPERMUTE)



BWT: Reverse Transform(UNPERMUTE)



BWT: Reverse Transform(UNPERMUTE)