Nucleics

CounterTrace™

Users manual

Version 1.3.0

February 2004

Table of contents

Overview

Current state of the art capillary DNA sequencers can produce trace reads that exceed 1000 nucleotides. Unfortunately, long trace reads typically suffer from excessive peak broadening and poor resolution in the later part of the trace. These latter regions are associated with low quality and high error rates.

CounterTrace uses a novel approach to increase sequencing data quality whereby a labelled DNA standard is introduced post-cycling into the reaction. The associated software extracts the raw data from the resulting trace files, identifies the location of the standard peaks, removes the standard peaks, reprocesses the trace data utilising the information provided by the standard peak locations, and base calls the trace using the inbuilt CounterTrace base caller. Alternatively, the resulting data can be base and quality called using phred or other external base calling software. CounterTrace is compatible with all common sequencing reagents and reaction conditions.

An example ABI 3730 trace from a 50cm capillary array before and after CounterTrace reprocessing is illustrated in Figure 1.
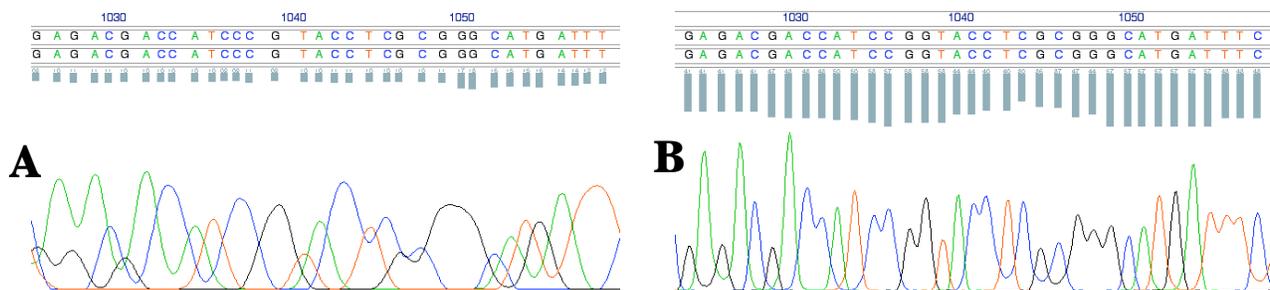

 Figure 1. Phred derived base and quality calls from an ABI3730 sequence trace. **A.** Before CounterTrace reprocessing. **B.** After CounterTrace reprocessing.

CounterTrace™ system components

CounterTrace Loading Additive

The *CounterTrace Loading Additive* (brown tube) should be **stored in the dark at -20°C** and is stable for at least 6 months when stored undiluted.

CounterTrace Software

The CounterTrace software is provided as command line binary files for use with Windows2000, Windows XP and x86 Linux. Versions for use with other operation systems can be supplied upon request.

CounterTrace Loading Additive

The *CounterTrace Loading Additive* is supplied as a 1000x stock in 20mM NaOH. The additive must be diluted to a 1–5x concentration in the loading buffer appropriate for the sequencer being used (i.e. water, 50µM EDTA, or formamide) before adding to the cleaned DNA sequencing samples. For example, if you want to resuspend 1000 sequencing reactions each in 10 µl, a 1x solution from 10 µl of 1000x CounterTrace stock and 9.9 ml of loading buffer should be made. It is critical that the additive is not diluted further than 1x. **The use of greater dilutions than 1x (e.g. 0.5x) may prevent the CounterTrace software from being able to detect the standard peak locations and hence prevent reprocessing of the trace file.**

If desired, a bulk mix of the loading additive can be made, however, it is recommended that bulk mixes are not stored for longer than 12 hours as the *CounterTrace Loading Additive* is significantly more light sensitive in the dilute form.

CounterTrace Software

1.1.   Overview

The CounterTrace software processes ABI format trace files for base and quality scoring using external software such as phred. Alternatively, the inbuilt CounterTrace base caller can be utilised, however, this base caller does not currently provide quality scores.

The CounterTrace software can process either single trace files or a folder of trace files. The output for each file is written to the specified output folder. If a user specifies an input folder then all files found within that folder are processed without recursion. When processing folders the CounterTrace software provides an estimate of the time of completion.

If during processing the CounterTrace software encounters an ABI trace file that cannot be processed then the software will copy the input file to the output path by default. This copying includes any necessary file format conversion. A user can turn off this feature using the `-nc` option.

The CounterTrace software writes a logfile of the run process. The beginning and end of each run is marked in the log, and if the file already exists, then the new run messages are appended to the existing log file. By default, this logfile is named `ctlog.txt`. A user can change this name using the `-l` option. The `-nl` option prevents the CounterTrace software from writing to the logfile. A summary of the processing is written to the log file detailing pass rate and reasons for why traces could not be processed (see the *CounterTrace software errors* section).

The default file output is ab1. The user can select between the ab1 and scf v3.1 file formats with the `-t` option. Unless storage space is limited, it is recommended that the ab1 output format is selected, as this format allows subsequent reprocessing of the trace data.

1.2.   Spectral Calibration

CounterTrace requires spectral calibration of the *CounterTrace Loading Additive* (standard) for each individual sequencing instrument used. This calibration must be performed before the first use of CounterTrace, or whenever an ABI3730 spectral calibration is recommended (please see the *Applied Biosystems 3730/3730xl DNA Analyzers Users Guide* for further information).

To perform the spectral calibration, dilute 1µl of the 1000x *CounterTrace Loading Additive* in 999µl of loading buffer. Aliquot 10µl of this 1x stock into either one quadrant of a 384-well tray, or all wells of a 96-well tray. Load this tray on the desired machine and run as per standard sequencing conditions.

This process should generate a folder of 96 standard-only traces (blanks). Running the CounterTrace software on this folder of traces using the "`-mc`" option will allow the software to determine the machine spectral calibration data for the individual machine on which the standard-only data was collected. This calibration data is stored in the binary file `machines.ctc`. No output folder is required in this run mode.

Subsequently, when run in the normal processing mode, the CounterTrace software will refer to this spectral calibration file. If a machine is used for which a CounterTrace spectral calibration has not been performed, then the CounterTrace software will not process traces produced on that machine until the above outlined spectral calibration is performed.

The location for storage and retrieval of calibration data is determined by the following order of precedence:

1. The path defined with the command line option " `-mp [path]` ".

2. The environment variable `COUNTERTRACE_DIR`.

3. The folder path of the CounterTrace executable.

When calibrating a machine, CounterTrace will require write access to this path. During normal processing CounterTrace requires only read access to the `machines.ctc` file.

CounterTrace software errors

The following error messages may occur when attempting to process trace files. The various possible causes of these errors are described below in alphabetical order together with possible solutions.

Can't create file

This error may occur if you do not have write permission to the selected output folder or if the hard disk is full. If the hard disk is near full please remove excess files and restart CounterTrace.

Can't find standard

This error indicates that CounterTrace cannot detect the standard peak location in the trace. This error is caused by either noisy trace data (i.e. multiple simultaneous peaks) or insufficient standard being added to the sequencing sample. Examine the raw data in the input trace file using an appropriate trace viewing application. If the trace appears to contain multiple simultaneous peaks (e.g. caused by mixed template) then remove the trace from the process list. If the trace does not appear to contain multiple simultaneous peaks then use a higher concentration of the CounterTrace Loading Additive (i.e. 1.5x or 2x). DO NOT USE DILUTIONS OF THE COUNTERTRACE STANDARD BELOW 1x.

Can't open file

This error may occur if the file does not exists or if the file has incorrect access permission settings. Check if file exists and that you read permission to the input folder.

Can't read file

This error may occur if an open file was deleted or moved while being used by CounterTrace. Do not delete or move trace files after starting CounterTrace.

Can't write file

This error may occur if you do not have write permission to the selected output folder or if the hard disk is full. Please remove excess files and restart CounterTrace and check if you have write permission to the folder selected.

Invalid file

This error may occur if the trace file is corrupt. Remove trace from process list.

Mixed or noisy data

This error may occur when processing traces containing mixed template. Examine the input trace file using an appropriate trace viewing application. If the trace appears to contain multiple simultaneous peaks (e.g. mixed template) then remove the trace from process list.

No bases in file

This error may occur if the trace file is corrupt. Remove trace from process list.

No base order in file

This error may occur if the trace file is corrupt. Remove trace from process list.

No calibration data for this machine

CounterTrace can only process traces from sequencing machines for which it has a calculated CounterTrace spectral calibration file. This error may occur when the CounterTrace spectral calibration has not been performed for the sequencer that produced the trace, or the spectral calibration data file cannot be found. Perform a CounterTrace spectral calibration for the respective sequencer.

Not enough memory

This error may occur if too many applications are running. Close nonessential applications and restart CounterTrace.

No machine name

This error may occur if the trace file is corrupt or if the sequencer has not been named. Remove trace from process list or see ABI sequencer user manual.

No peak locations in file

This error may occur if the trace file is corrupt. Remove trace from process list.

No processed data in file

This error may occur if the trace file is corrupt. Remove trace from process list.

No raw data in file

This error may occur if the trace file is corrupt. Remove trace from process list.

Processing took too long

This error occurs if the trace cannot be processed within one minute. This error may occur if using CounterTrace on slow computers (800Mhz PentiumIII or less), or by a program malfunction. Use a faster computer for processing. If this error occurs when using a fast computer please contact Nucleics.

Process was aborted

This error will occur if a CounterTrace sub-process is stopped explicitly. Do not stop the CounterTrace sub-processes.

Program error

This error indicates a program malfunction. Please contact Nucleics.

Spacing model

This error indicates an error in identifying the correct peak spacing. This maybe caused by mixed or noisy data or other problems with the trace data. If this error occurs with a high percentage of your traces please contact Nucleics.

Trace end too weak

This error indicates that the trace file does not contain sufficient peak signal after base 700 for 36cm arrays and after base 800 for 50cm arrays. CounterTrace can only process traces that are limited by the resolution of the sequencing instrument, not traces that are limited by signal intensity. Increase sample injection times and/or modify sample preparation conditions to ensure that sufficient peak signal is available after base 700–800.

Traces from multiple machines

This error occurs when performing a CounterTrace spectral calibration. Spectral calibration requires that a folder containing trace files for calibration are all generated from the same sequencing machine. Only use traces from the one machine when generating a CounterTrace spectral calibration.

Unknown machine type

This error may occur if processing a trace file produced by a sequencing machine type that CounterTrace does not support or if the trace file is corrupt. Remove the trace from process list.

Weak or failed capillary

This error indicates that the trace contains very weak or no peak signal. Traces giving this error have either completely failed or have very weak signal. Traces that give this error rarely have read lengths over 100 bases.

Command line switches and parameters

The CounterTrace software has the following command line parameters and options:

1.3.   help   `-h`

Lists the help information for the CounterTrace software

1.4.   version   `-V`

Lists the version number of the CounterTrace software

1.5.   input   `-i [string]`

Specifies the input file or folder. If a folder is specified all files within that folder are processed. No recursion is performed on folders.

1.6.   output   `-o [string]`

Specifies the output path (folder) for the processed trace files. For each input file a processed trace file will be written to this output folder. Output file names are based on the input file names. If the user has specified an alternate output file format then the trailing suffix is altered or added to reflect the new format.

The output folder must exist before processing and it cannot be the input folder. Files with the same name in the output folder are overwritten.

If the CounterTrace software cannot process a trace file then the default behaviour is for the input trace to be copied (with any necessary file format conversion) to the output path. This feature can be turned off with the option `-nc`. Thus by default, the output folder should always contain the same number of traces as were in the input path unless user specified.

1.7.   output file format   `-t [abi|scf]`

Specifies the output file format. The two choices are ab1 or scf v3.1. Output files will reflect the chosen file format by the suffix `.ab1` or `.scf.`

The default is ab1.

1.8.   log name   `-l [string]`

Specifies the name of the processing status logfile. This log is always appended to rather than written over.

The default log file name is `ct.log`, however, both the log name and path can be user defined.

e.g.   unix: `-l /usr/mylogs/ct.log`

windows: `-l c:\my documents\ct.log`

1.9.   no copy   `-nc`

Prevents copying of non-processed input files to the output path.

Default is OFF.

1.10.   no log   `-nl`

Prevents the processing status being written to the log file.

Default is OFF.

1.11.   spectral calibration   `-mc -i [string]`

Generates a machine-specific spectral calibration file for CounterTrace that aids in the accurate subtraction of the CounterTrace standard. The input must be a folder containing a set of standard-only ab1 trace files from a single machine. This calibration data is stored in the binary file `machines.ctc`. No output folder is allowed in this run mode.

1.12.   multi-processor   `-j [integer]`

Allows the CounterTrace software to run multiple sub-processes. This allows multiple traces from an input folder to be process in parallel. This setting is of utility on multi-processor machines or cluster arrays. On single processor machines it is not recommended that more than one sub-process is used.

Default is ONE.

Technical Support

Nucleics provides expert technical support for the CounterTrace™ system.

Please address any questions or comments to:

countertrace@nucleics.com