

Ridom TraceEdit: a DNA trace editor and viewer

Jörg Rothgänger¹, Markus Weniger¹, Thomas Weniger², Alexander Mellmann³ and Dag Harmsen^{1,2,*}¹Ridom GmbH, 97082 Würzburg, Germany, ²Department of Periodontology, University Hospital Münster, 48149 Münster, Germany and ³Institute for Hygiene, University Hospital Münster, 48149 Münster, Germany**ABSTRACT**

Summary: Ridom TraceEdit is a cross-platform graphical DNA trace viewer and editor. TraceEdit displays the chromatogram files from Applied Biosystems automated sequencers and files in the Staden SCF format. Incorrect base calls can be edited and saved.

Availability: TraceEdit is freely available via the Internet at <http://www.ridom.de/traceedit/>. TraceEdit has been designed to operate on Windows, MacOS X and UNIX platforms.

Contact: info@ridom.de

DNA sequencing has been the standard to which other types of DNA testing has been compared. Major advances in DNA sequencing include the development of automated sequencers, discovery of fluorescent terminator chemistry, and cycle sequencing. These developments have facilitated sequencing and, therefore, it is more widely used. Currently, sequencing is being used to identify microbial drug resistance mutations, cancer predisposition, somatic mutations, and genetic diseases. With the sequencing of the human genome and the “new era of molecular medicine”, one can only expect the use of DNA sequencing even to increase (Murphy and Eshleman, 2002).

Therefore, an increasing number of people need to check the evidence for individual DNA sequences by inspecting the chromatograms (more commonly known as trace files) from which the base calls were deduced. Trace files are stored in proprietary formats, such as those of ABI, or public formats such as SCF (Dear and Staden, 1992). These files contain the trace amplitudes, the base calls, their quality values, and textual data about the particular sequencing experiment such as its chemistry, machine type and operating conditions, etc. The first non-commercial trace editor and viewer ted (Gleeson and Hillier, 1991) was later replaced by the software trev (Bonfield *et al.*, 2002). Trev has recently become an open source project and, therefore, is no longer advanced and supported in an institutional framework. Furthermore, it is somewhat cumbersome to install (requires Tcl). Therefore, new trace editor software - Ridom TraceEdit - was designed and programmed. The platform-independent Java technology was used to provide distributions for various operating systems (e.g., Windows, MacOS X and Linux) and hardware architectures.

Ridom TraceEdit can read ABI and SCF trace files and display their contents. The program can be used to locate and delete low

quality data at both ends of the sequence (trimming by called Ns or low base quality values if available). It also enables the user to edit the base calls and save them in SCF file format. It includes from top to down the name of and additional information on the inspected sequence, the menu bar, the file list pane with tool bar, the trace editor pane with tool bar, and a status bar. A typical display from TraceEdit is shown in Figure 1. In the file list pane a trace file can be chosen to be displayed in the editor pane. The trace editor pane shows the trace data, the quality values (if available), the original sequence, the edited sequence, and the base positions. Poor quality data at the start and end of the reading are denoted by gray shading. A separate text window (not shown) is used to display the textual data contained in the file.

The trace data can be scrolled using the scrollbar beneath the trace editor pane. The trace can be magnified in horizontal and vertical direction. The display of the original sequence and the quality values (different formats are available) can either be switched on or off. It is possible to open several traces at once and shortcuts allow for quick navigation between traces. TraceEdit either produces PDF files of the traces or sends them directly to the printer. The user can define, store, and search for nucleotide patterns that may include regular expressions (Figure 1). The edited sequence can be copied to the clipboard or exported to a file. Alternatively, a NCBI nucleotide database BLAST search may be initiated with the edited sequence directly from TraceEdit (Wheeler *et al.*, 2005). Sequences from the NCBI nucleotide database can be imported directly into the software by stating the database accession numbers. The software behavior may be user-defined in several aspects (e.g., editor preferences, trimming options, dialog options, or automatic search for updates) by using the Options/Preferences menu entry (for further details and shortcuts it is referred to the accompanying manual).

TraceEdit includes a feedback dialog to provide the user with a convenient way of sending messages such as feature requests and bug reports directly to the developers. Finally, a download option for online updates is integrated into the software.

Several other freeware trace editors are meanwhile available, i.e., in chronological order of release: EditView (Applied Biosystems, Forster City, CA; MacOS < X), Chromas Lite (Technelysium Pty Ltd, Australia; Windows), 4Peaks (www.mekentosj.com, MacOS X), and FinchTV (Geospiza Inc, Seattle, WA; Windows, MacOS X, UNIX). The basic functions of these programs – including

*To whom correspondence should be addressed.

TraceEdit – are similar. However, there are marked differences regarding the number of supported platforms and some additional features. The latest developments TraceEdit and FinchTV support the broadest number of platforms. FinchTV features a unique wrapped multi-panel display of chromatograms, whereas TraceEdit offers automatic trimming options, seamless import of NCBI nucleotide database sequences, and easy handling of multiple files by using the file list pane functions. The source code of none of the above mentioned programs is free.

However, open source trace viewers (not editors) are available. All of them are based on Java, i.e., Jeremy Parsons DNA sequence chromatogram browser (Parsons *et al.*, 1999) originally used for Ensembl's trace archive, a trace viewer written by Kathryn Beal that is currently used by Ensembl (Hubbard *et al.*, 2005), and finally a Java Applet written by Andriy Klimenko and Sergiy Ponomarov for the retrieval and viewing of trace data from the NCBI Trace Archive. All three programs are primarily designed to interact as a client with a central server. Freeware trace editors like TraceEdit, therefore, currently offer greater flexibility and functionality on a user desktop as a stand-alone application.

Hubbard, T. *et al.* (2005) Ensembl 2005. *Nucleic Acids Res.*, **33**, D447-D453.

Murphy, K.M. and Eshleman, J.R. (2002) Simultaneous sequencing of multiple polymerase chain reaction products and combined polymerase chain reaction with cycle sequencing in single reactions. *Am. J. Pathol.*, **161**, 27-33.

Parsons, J.D. *et al.* (1999) DNA sequence chromatogram browsing using Java and COBRA. *Genome Res.*, **9**, 277-281.

Wheeler, D.L. *et al.* (2005) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.*, **33**, D39-D45.

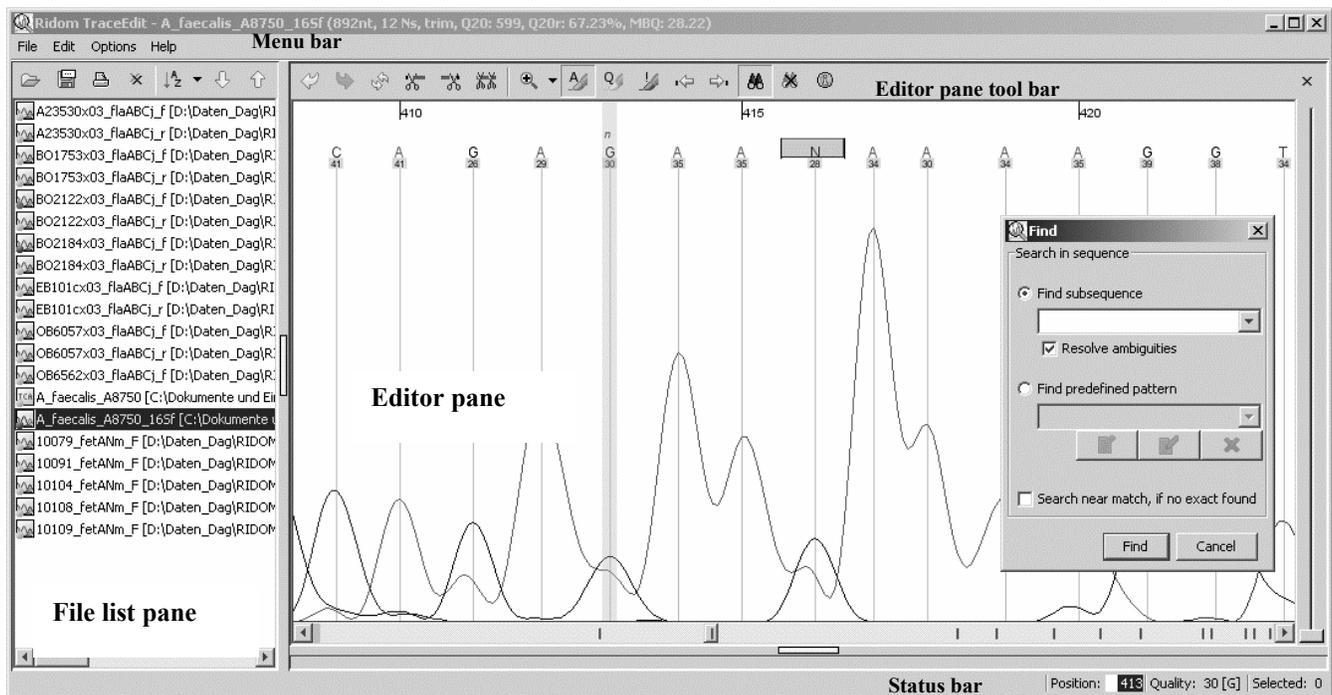


Fig. 1. Screen dump from Ridom TraceEdit (graphical user interface elements like the editor pane are marked by text). See the software manual and the text for further details.

REFERENCES

- Bonfield, J.K. *et al.* (2002) Trev: a DNA trace editor and viewer. *Bioinformatics*, **18**, 194-195.
- Dear, S. and Staden, R. (1992) A standard file format for data from DNA sequencing instruments. *DNA Sequence*, **3**, 107-110.
- Gleeson, T. and Hillier, L. (1991) A trace display and editing program for data from fluorescence based sequencing machines. *Nucleic Acids Res.*, **19**, 6481-6483.