Review

Loader Lite: A New Software Tool for the ABI PRISM® 3700 DNA Sequencer

BioTechniques 32:1366-1371 (June 2002)

G.B.I. Scott, D.L. Steffen, D. Edgar, J.T. Warren, C.L. Kovář, S.E. Scherer, P.H. Havlak, and R.A. Gibbs
Baylor College of Medicine,
Houston, TX, USA

have successfully used Loader Lite over the past six months to load approximately 7 million sequencing reactions and believe its utility and functionality will prove to be attractive to the wider sequencing community.

ABSTRACT

Here we describe the development of a novel software tool entitled Loader Lite that generates plate records or sample sheets for the ABI PRISM® 3700 DNA sequencer. The major advantage of this program is that it enables the ongoing operation of sequencing instruments without reference to external network(s). The autonomous operation of sequencing instruments is critical if sample throughput is to be maintained during periods of network outage. Loader Lite employs a deliberate strategy of inputting anonymous tray barcodes at run time. After sequencing, the barcodes are reconciled with relevant project details by reference to a database. This software takes advantage of barcode scanning technology by creating plate records directly on the local computer, serving an individual sequencer, immediately before importing and linking. This real-time synthesis of the plate records at the point of loading all but eliminates loading errors. Loader Lite is user-friendly, fully configurable, and permits the running of partial or full 384-well sample trays, using any standard combinations of run modules, dye sets, mobility files, analysis modules, etc. The 96-well format is not supported; however, this capability will appear in subsequent versions that are currently under development. This application is designed as an added value, adjunct program to the regular ABI PRISM 3700 Data Collection software. We

INTRODUCTION

In recent years, the ABI PRISM® 3700 DNA Sequencer (Applied Biosystems, Foster City, CA, USA) has become the preeminent tool of choice (3–6) for conducting high-throughput DNA sequencing operations. This instrument affords considerable utility, primarily by permitting the largely unattended, sequential, and continuous analysis of up to four 384-well microplates containing DNA samples (ABI PRISM 3700 DNA Sequencer User Manual). Applied Biosystems has distributed in excess of 1000 of the ABI PRISM3700 DNA sequencers. The degree of automation, reliability, and throughput conferred by clusters of these instruments has allowed the Human and other Genome Projects to progress rapidly toward completion.

One of the most problematic aspects of high-throughput, production-scale DNA sequencing is the management (i.e., tracking, storage, retrieval, quality control, statistical analysis, etc.) of the large data stream that inevitably results from sequencing activities. In response, all large genome centers have aggressively recruited increasingly large numbers of computer professionals, programmers, and bioinformaticians to address these complex issues.

A central operational issue addressed by all genome centers is the production

of plate records or sample sheets that are used to accurately record the identity of each sample tray/well that is analyzed on the ABI PRISM 3700 instruments. Traditionally, sample sheets have been generated directly using the Data Collection software that is provided with each sequencer or by using spreadsheet software tools such as Microsoft® Excel®. Both methods are relatively slow and involve considerable manual entry by the user, but they are flexible and suitable for use in small- and medium-scale sequencing environments. These methods are not amenable for use in large genome centers; thus, they have created in-house software solutions to automate the generation of ABI PRISM 3700 instrument sample sheets (6). Most programs automatically generate a sample sheet on typing in a tray identifier or scanning a barcode on the sample tray. Such programs typically query a database, and the sample sheets generated are usually stored on a centralized server or mainframe computer. When a tray is ready for loading, the sample sheet is imported from the server into the Data Collection software before linking and run commencement. This scheme suffers from several deficiencies, including the front-end reliance on a database that will recognize a typed or scanned plate identifier and subsequently generate a sample sheet. In turn, this generates a dependence on database availability and a network that enables access. However, even the most robust databases and networks are somewhat unreliable and need to be periodically taken off-line for servicing or upgrade, resulting in disruptions to sequencing operations. The other problem

BioComputing/BioInformatics>>>>:

encountered when using the server/network methodology relates to tracking issues. Conventionally, plate records are created en masse using a dedicated remote terminal and then are imported piecemeal from the centralized server onto the sequencing instruments. This type of approach relies on stringent attention to detail during the loading, importing, and linking phases to ensure that mix-ups do not occur. The most typical mistake involves operator error in which the plate record stored in the Data Collection software and the barcode affixed to the corresponding physical plate do not match.

Here we describe a novel software solution for generating plate records that has several innovative features and largely overcomes these two major problems. We call this program Loader Lite.

DISCUSSION

The concept behind Loader Lite is that plate records or sample sheets are generated locally on the computer connected to each ABI PRISM3700 instrument, thus obviating the need to conduct sequencing operations with constant reference to external networks or servers. Moreover, although Loader Lite is a stand-alone application, it is designed as an adjunct program that both interfaces seamlessly with and adds value to the standard Data Collection software. Additionally, Loader Lite uses barcode scanners so that the hand entry of data is eliminated. Finally and most significantly, Loader Lite provides for the scanning of anonymous sample tray barcodes at run time, with reconciliation between barcode and project-identifying parameters occurring downstream of the run.

Loader Lite is written in Visual Basic code and may be run on any computer running the Windows NT® operating system. This program has a small footprint, with the executable file being only 336 kb, and it requires minimal CPU and memory resources. Loader Lite is typically quiescent during the execution of the Data Collection software and is only used before commencing sequencing runs.

A key advantage of Loader Lite lies in its simplicity and ease of use. The

user places any prepared 384-well sample tray(s) on any available plate positions (i.e., A–D) on each ABI PRISM 3700 sequencer. Loader Lite does not currently support the 96-well plate format, but this feature is under development and will be added in the future. The Loader Lite software is opened by double-clicking on a shortcut icon that is created on the desktop of each computer at the time of initial software installation. The user is first prompted to scan their identification badge or, alternatively, type in an identification number that is unique to the organization or company. The input of a valid (currently six-digit) identification number will display a subsequent form to the user that allows for the synthesis of up to four plate records (Figure 1). The user then employs a barcode scanner to scan a single sample tray barcode, immediately followed by the unique barcode for the particular plate deck location (A, B, C, or D) on the sequencer. Note that the barcodes denoting the various plate positions are affixed to the plate deck of the sequencer. The graphical user interface is updated in real time, according to the inputs received from scanning [i.e., the relevant tray identification numbers are displayed in the appropriate text boxes below the plate position depictions (Figure 1)]. Finally, the software offers further flexibility by invoking the Customize... button, which enables the

user to run partial 384-well trays; that is, one or more 96-well quadrants of a sample tray may be selected or deselected as desired (Figure 2).

Once the user has scanned in the entire batch of sample trays and their concomitant plate deck positions, a final barcode that is affixed to the lid of the ABI PRISM 3700 instrument is scanned to signify that the synthesis of plate records is complete. Alternatively, clicking the Generate Data button also concludes the scanning procedure. If a mistake is made during the scanning process, then the procedure may be aborted and the software reset by scanning a “Clear” barcode that is affixed to the oven door on the ABI PRISM3700 instrument or by clicking the Clear All button on the main user interface. After the user has satisfactorily completed the scanning process, they exit the program. Accordingly, Loader Lite does not typically run in parallel with the ABI PRISM Data Collection software. If a user fails to exit Loader Lite after the generation of one or more plate records, then the program will run benignly in the background.

On the completion of the above steps, Loader Lite creates one or more sample sheets within the Sample Sheets directory of the C drive of the local computer. This directory has subdirectories that correspond to calendar days, and plate records are placed into the subdirectory corresponding to the cur-

Figure 1. View of the Main or Sequencer Layout screen. Screen shows sample trays that were scanned into plate positions A, B, and C. Note that plate position D is vacant.

BioComputing/BioInformatics>>>>:

rent date. If the Date folder does not exist at the time of scanning, then Loader Lite will automatically create a Date folder within the Sample Sheets directory by referencing the Date/Time setting of the local computer.

The plate records that are generated by Loader Lite have all the appropriate fields filled in, including Sample Name, Dye Set, Mobility File, Project Name, Run Module(s), and Analysis Module. Note that Loader Lite creates plate records in which certain fields are necessarily fixed and identical for each discrete 96-well load. In this regard, the application is restricted to generating plate records with one type of chemistry and, typically, one or two primers per sample plate. Loader Lite is efficient in recognizing and processing a small range of standard sample types by referencing specific (optional) characters embedded within the anonymous barcode. In this way, Loader Lite synthesizes a correctly parameterized plate record that is appropriate to the generic sample type. Loader Lite was not developed for the detailed naming of subclones (with the concomitant input of varying run and analysis parameters) typically used in core laboratories or to process finishing reactions.

At this point, the user will employ the ABI PRISM Data Collection software to initiate the importing and linking of the desired plate records by clicking on the Import button available on the Plate Setup tab. Clicking the Import button allows the user to navigate to the appropriate subdirectory Date folder and locate the desired plate records. The user selects each plate record in turn and then either imports and/or imports and links in one step, as desired. The manual selection of the appropriate plate record from the Sample Sheets Folder is required, and there is still a small but finite possibility of user error (i.e., mis-linking) during this last step. Only one instance of mis-linking has been seen during our six months of experience using Loader Lite, which represents an error rate at least one order of magnitude less than experienced before the implementation of Loader Lite. To amend the plate record fields, the user may click on the Open button within the Data Collection software and make desired changes.

The Loader Lite software references an initialization (.ini) file that provides setup information pertinent to each discrete sequencer. This .ini file is small, only 7 kb in size. Software that generates plate records needs to be flexible and allow for standard variations in configuration between sequencers, such as the type of matrix (e.g., POP 5 or POP 6) run or the dye set used (e.g., BODIPY® and ABI PRISM BigDye™ versions 1 and 3). The initialization file allows for full editorial control over the generation of the plate record by facilitating the amendment of the unique sequencer identification number, along with the entries that will be entered into the fields within the plate record—such as the mobility file(s), analysis, and run module(s). Loader Lite is thus fully configurable. Finally, the latest version of the program (version 1.5) allows for changes to be made to the initialization file via a user interface, rather than by directly locating, accessing, and amending the .ini text file (Figure 3).

High-throughput DNA sequencing operations demand the efficient management, transfer, analysis, and archiving of completed run files or “traces”. This data is typically transferred automatically off the local computer(s) that

serve the ABI PRISM 3700 sequencers and onto UNIX-based systems, where the downstream analysis is performed using such programs as Phred (1,2) and Phrap. Loader Lite uses anonymous or coded identification tray barcodes that do not provide *a priori* the parameters that identify specific projects. Also, the tray barcode does not directly reference any database at the time of scanning. Accordingly, the sample tray barcode must be reconciled with the main (project) database, following the extraction and transfer of the “abi” data files onto the UNIX system. In this way, the anonymous barcode and project-identifying parameters are mapped/linked downstream of the sequencing run. This reconciliation task is accomplished using additional interface software that continuously interrogates the Oracle® project database (Oracle, Redwood Shores, CA, USA) and merges projects with matching sequencer traces.

CONCLUSION

We have described a new software tool for the generation of plate records in situ on the computer serving each ABI PRISM 3700 DNA sequencer. This

Figure 2. View of the Standard Customization screen. Individual 96-well quadrants of a 384-well plate may be selected or deselected as required by checking or unchecking the boxes adjacent to the listed run modules. The graphical color-coded depiction of the sample tray is automatically updated according to the selections made. Advanced customization features are accessed by clicking on the Advanced... button.

Figure 3. View of the Customization screen with Advanced Features displayed. Advanced Features allow users to modify various run parameters such as the Mobility File, Dye Set, Analysis Module, Sequencer Name, Number, etc., for the current run. The option to permanently save the new settings to the configuration (i.e., .ini) file is available using the command buttons, Update Run modules or Update Sequencer Info.

software performs as an adjunct module while enhancing the utility and adding value to the standard ABI PRISM Data Collection software.

The salient advantages of the Loader Lite software are its ability to run without reference to external network(s), ease of use, and the reduced propensity for mix-ups in tray importing and linking. The local or “un-networked” nature of Loader Lite obviates the need to maintain 100% network availability. This stand-alone feature is advantageous for medium- and large-scale genome centers that both operate complex local area networks and run samples continuously (i.e., nights and weekends). The second important feature of Loader Lite is its ease of use by incorporating barcode scanning technology, thereby eliminating any need for the manual entry of data, except in a few special cases. Finally, sample sheets are generated in situ on the local computer serving each instrument, following the positioning of the sample plates on the plate deck of the ABI PRISM 3700 sequencer. This simple process involves successive scanning of a series of barcodes pertaining to the user identification number, the tray identification number, the sequencer/

plate position, and a “process complete” or “redo” barcode. Because the plate position and sample tray barcodes are scanned after the tray has been locked into position, the possibility of mix-ups caused by sample identification is all but eliminated.

Loader Lite has been used successfully in our high-throughput DNA sequencing operations for the past six months, over which time roughly 7 million DNA sequencing reactions have been analyzed. We contend that this software tool will be of considerable utility for all genome centers using the ABI PRISM3700 sequencing platform. Copies of the Loader Lite software are available at no charge to nonprofit institutions from the corresponding author.

REFERENCES

1. **Ewing, B. and P. Green.** 1998. Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res.* 8:186-194.
2. **Ewing, B., L. Hillier, M.C. Wendt, and P. Green.** 1998. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res.* 8:175-185.
3. **Friend, T.** Genome mappers raced to middle ground. *USA Today* 2000 June 27; Sect D:10.
4. **Golden, F. and M.D. Lemonick.** 2000. The race is over. *Time* 156:18-23.

5. **Marziali, A. and M. Akeson.** 2001. New DNA sequencing methods. *Annu. Rev. Biomed. Eng.* 3:195-223.
6. **Venter, J.C., M.D. Adams, E.W. Myers, P.W. Li, R.J. Mural, G.G. Sutton, H.O. Smith, and M. Yandall.** 2001. The sequence of the human genome. *Science* 291:1304-1351.

Received 12 October 2001; accepted 9 April 2002.

Address correspondence to:

Dr. Graham B.I. Scott
Baylor College of Medicine
Human Genome Sequencing Center
One Baylor Plaza
Houston, TX 77030, USA
e-mail: gscott@bcm.tmc.edu

**For reprints of this or
any other article, contact
Reprints@BioTechniques.com**