Sequencing Difficult DNA Templates: Comparing Capillary Electrophoresis Systems

Introduction

For many researchers, examining difficult DNA templates is an essential but complex part of their sequencing routine. Template regions with above-average sequencing complexity are not uncommon. Most complex templates contain a feature, such as a GC-rich region or a hairpin loop, which may inhibit the activity of DNA polymerase. Long homopolymer strings can also reduce the effectiveness of certain sequencing chemistries. Difficult templates are expensive and time-consuming, as researchers try troubleshooting strategies such as experimenting with different reaction conditions, different sequencing chemistries, sequencing the opposite strand, and additional manual trimming and editing.

This fact sheet compares the ability of the Applied Biosystems 3130 and 3130*xl* Genetic Analyzers and BigDye® Terminator sequencing chemistries, with that of a competitive system to handle difficult DNA template sequencing. These templates include repeat regions, GC-rich templates and regions, and BAC end-sequencing.

G/A repeat data sequenced on the 3130*xl* system

G/A repeat data sequenced on a competitive system

T/A repeat data sequenced on the 3130*xl* system

T/A repeat data sequenced on a competitive system

Figure 1. Data generated from a plasmid containing G/A or T/A repeats using the 3130*xl* Genetic Analyzer, BigDye® Terminator v3.1 chemistry, and Sequencing Analysis Software v5.2 versus a competitive system. The 3130 Series Systems provide even peak heights and resolution across both the G/A and T/A repeat regions using standard protocols.

Ideally, a system consisting of a sequencing chemistry, analysis software, and capillary-based genetic analysis instrument should be able to produce high-quality data from difficult templates with little or no modification to protocol. When evaluating a capillary electrophoresis (CE) system, these factors are of prime importance (Table 1).

Repeat Regions

The incorporation efficiency of a sequencing reaction is greatly affected by the template base composition. The robustness of a sequencing system reveals itself in quality sequence reads, read length, signal-to-noise ratio, and basecalling accuracy. The difference between the Applied Biosystems 3130 Series Systems and a competitive system is readily apparent when sequencing templates that contain difficult base context, such as repeat regions.

To evaluate the performance of the 3130 Series System and BigDye® Terminator v3.1 chemistry with respect to repeat regions, scientists at Applied Biosystems sequenced and analyzed a template containing a G/A

Table 1. Features to Consider in a CE System

CE instruments should provide:	• High sensitivity to dye-labeled sequencing reaction products* • Low background fluorescence* • Precise control of electrophoretic separation conditions*
Sequencing chemistries should provide:	• Extension across difficult regions • Uniform dye incorporation • Robust signal
Analysis software should provide:	• Accurate basecalling results with quality values • Accurate variant detection, including mixed bases and heterozygous insertions and deletions • Easy quality control • Exportable and comprehensive results reporting

These features enable the analysis software to interpret data accurately with the highest quality value (QV) estimations.

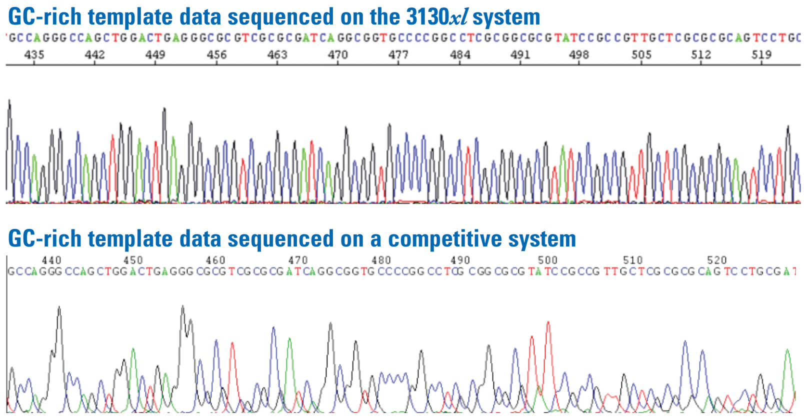


Figure 2. Comparison of two templates with high GC content, sequenced on the 3130*xl* Genetic Analyzer and a competitive CE instrument. The 3130*xl* system provides well-resolved peaks without protocol modifications.

and T/A repeat region. The BigDye Terminator chemistry was able to sequence across both repeat motifs easily, producing uniform peak heights, while the 3130*xl* system produced high-resolution peaks, especially in the homopolymer stretches. Algorithmic analysis using Sequencing Analysis Software v5.2 and KB™ Basecaller v1.2 provides high-quality basecalling, which results in efficient completion of sequencing projects that include difficult templates (Figure 1).

GC-Rich Templates

GC-rich templates have unique secondary structures and melting temperatures. These characteristics occur when a greater proportion of nucleotides throughout the template have three hydrogen bonds, which maintain a stronger association between the double-stranded DNA. DNA sequences with a high GC content are relatively more difficult to denature and process, and may lead to a more challenging sequencing reaction.

Applied Biosystems scientists tested the performance of the 3130*xl* system, using the standard protocol for BigDye Terminator v3.1 chemistry on a GC-rich DNA template (Figure 2). The 3130*xl* system and BigDye Terminator chemistry processed the template easily, providing characteristically even and well-resolved peaks without protocol modifications. Furthermore, the new, higher temperature (60°C) 3130 POP-7™ Polymer run modules and the detection cell heater provide better thermal control and help reduce GC-rich secondary structures, leading to improved CE analysis.

Our competitor's system analyzed the same GC-rich template, using the manufacturer's recommended protocol. The result was data with broadly overlapping, uneven peaks, which could result in basecalling errors. Researchers who use other systems may have to invest additional time and money optimizing their protocols to achieve high-quality sequences from GC-rich templates.

GC-Rich Regions

Many DNA templates not considered GC-rich contain isolated GC-rich motifs that can present a challenge for various CE systems. In particular, progressive loss of DNA polymerase activity can result in low signal strength and reduced data quality for the template region immediately downstream of the GC-rich region.

The performance of the 3130 Series Systems running BigDye Terminator v3.1 chemistry was tested using a template bearing a homopolymer G stretch of DNA (Figure 3). Signal strength and peak-height uniformity were nearly the same downstream as upstream of the polyG segment. Additionally, all peaks in the area are well resolved. In contrast, the competitive system produced data with results that were minimally acceptable and possibly required additional sequencing strategies to achieve data quality adequate for successful project analysis and completion.

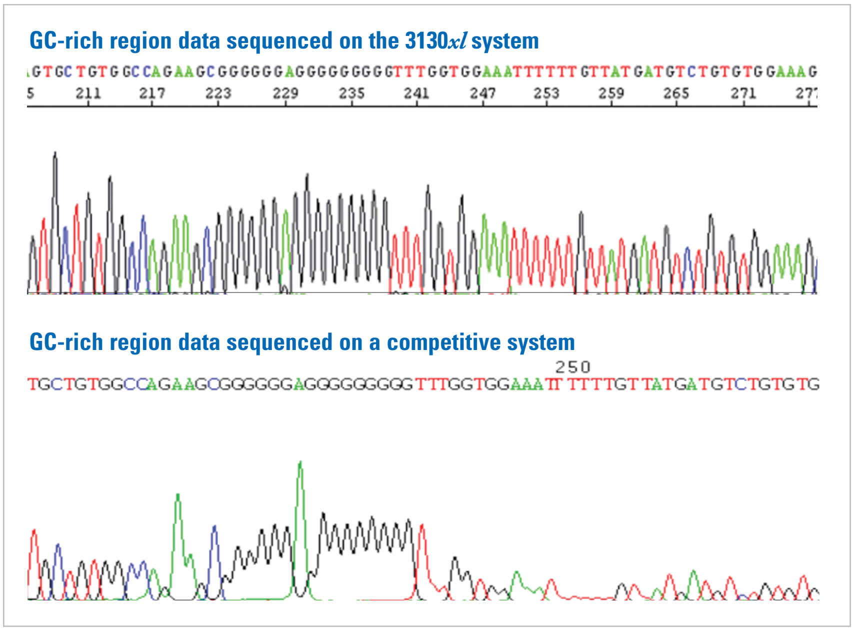
BAC End-Sequencing

Characterizing the ends of bacterial artificial chromosome (BAC) clones in genome-mapping studies entails difficulties resulting from the large size of the BAC clone.

Standard protocols for BAC end-sequencing require 2.5–3.0 µg of purified BAC DNA template per reaction. Preparing a sufficient quantity of template DNA for these protocols requires the growth of large volumes of BAC clone for subsequent DNA purification, which adds substantially to the time and expense required for the completion of these sequencing projects.

The robustness of Applied Biosystems BigDye® Terminator Sequencing Kits, coupled with the high sensitivity, improved thermal control, and resolving power of the 3130 Series Systems and 3130 POP-7 Polymer, lets researchers determine BAC end-sequences using less BAC DNA.

The standard DNA sequencing chemistry protocols developed by Applied Biosystems for this type of template can provide high-quality BAC end-sequence data on the 3130 Series Systems, with as little as 300 ng of starting template. The analyzed trace retains an even peak signal throughout the sequence read, characteristic of the balanced incorporation pattern common to BigDye kit chemistry. (Figure 4).


GC-rich region data sequenced on the 3130*xl* system

GC-rich region data sequenced on a competitive system

Figure 3. Template containing a homopolymer G stretch, analyzed on the 3130*xl* Genetic Analyzer and a competitive instrument. Data generated on the 3130*xl* system exhibits uniform signal strength and peak heights on either side of the poly-G region.


BAC end data sequenced on the 3130*xl* system

Figure 4. Analysis of BAC clone DNA sequenced on a 3130*xl* Genetic Analyzer using the 36 cm Rapid Sequencing POP-7 polymer run module. The 300 ng of starting template was sequenced according to the standard BigDye Terminator v3.1 protocol, except for the following substitutions, which were made to achieve the maximum quantity of high-quality data: primer amount: 12.0 pmol; sequencing reaction: 50 cycles.

Conclusion

The 3130 series instruments, BigDye Terminator chemistries, and sequencing analysis software, unlike competitive systems, provide high-quality data from difficult DNA regions without a substantial loss of accuracy or read length, and minimizes costly modifications to sequencing reactions or protocols.

iScience. To better understand the complex interaction of biological systems, life scientists are developing revolutionary approaches to discovery that unite technology, informatics, and traditional laboratory research. In partnership with our customers, Applied Biosystems provides the innovative products, services, and knowledge resources that make this new, **Integrated Science** possible.

Worldwide Sales Offices
Applied Biosystems vast distribution and service network, composed of highly trained support and applications personnel, reaches 150 countries on six continents. For international office locations, please call the division headquarters or refer to our Web site at **www.appliedbiosystems.com**

Applera is committed to providing the world's leading technology and information for life scientists. Applera Corporation consists of the Applied Biosystems and Celera Genomics businesses.

Headquarters
850 Lincoln Centre Drive
Foster City, CA 94404 USA
Phone: 650.638.5800
Toll Free: 800.345.5224
Fax: 650.638.5884

For Research Use Only.
Not for use in diagnostic procedures.

NOTICE TO PURCHASER:
Please refer to the Applied Biosystems 3130/3130*xl* Genetic Analyzer and GeneMapper software user's manual for limited label license or disclaimer information.

The Applied Biosystems 3130 and 3130*xl* Genetic Analyzers includes patented technology licensed from Hitachi, Ltd. as part of a strategic partnership between Applied Biosystems and Hitachi, Ltd., as well as patented technology of Applied Biosystems.

Applied Biosystems and BigDye are registered trademarks and AB (Design), Applera, iScience, iScience (Design), and POP-7 are trademarks of Applera Corporation or its subsidiaries in the US and/or certain other countries.

©2004. Applied Biosystems. All Rights Reserved.

Printed in the USA, 11/04,
P+s, Publication 106MI62-01

Applied Biosystems