The Use of Novel Enzymes for Sequencing with BigDye™ Terminators and a New, More Spectrally Resolved Dye Set.

Sandra L. Spurgeon[1], Diane R. Bond[1], James Rozzelle[1], Curtis Bloom[2], John Brandis[1], Applied Biosystems[1], 850 Lincoln Centre Drive, Foster City, California 94404, and the Department of Biochemistry[2], California Institute of Technology, Pasadena, CA 91125.

12th International Genome Sequencing and Analysis Conference

Miami Beach
September 12-15, 2000

Abstract

The Use of Novel Enzymes for Sequencing with BigDye™ Terminators and a New, More Spectrally Resolved Dye Set. Sandra L. Spurgeon[1], Diane R. Bond[1], James Rozzelle[1], Curtis Bloom[2], John Brandis[1], Applied Biosystems, 850 Lincoln Centre Drive, Foster City, California 94404 [1], and the Department of Biochemistry, California Institute of Technology, Pasadena, CA 91125[2]

The use of the BigDye™ terminator chemistry results in a 4-5-fold improvement in the signal-to-noise ratio and more consistent peak heights when compared to those obtained with rhodamine and drhodamine terminators. The enzyme used in this chemistry is AmpliTaq® DNA Polymerase FS, which has a very low level of 5'→3' nuclease activity and readily incorporates dideoxynucleotides. We have also used this enzyme to sequence with a new, more spectrally resolved dye set (see reference number 13). Although AmpliTaq® DNA Polymerase FS has excellent performance with both these dye sets, we have been interested in how other enzymes such as *Tth* DNA Polymerase compare. *Tth* DNA Polymerase is similar to AmpliTaq® DNA Polymerase but shows differences in its salt tolerance, DNA binding and other properties. We have also looked at the use of certain variants of AmpliTaq® DNA Polymerase FS that show altered ability to incorporate the BigDye™ terminators. Some of these variants show significantly improved incorporation of the dideoxyU terminator and are of interest for that reason. We have chosen one of these variants for further testing and compared it and a modified form of *Tth* DNA Polymerase to AmpliTaq® DNA Polymerase FS for use in DNA sequencing with BigDye™ terminators.

Introduction

The development of the BigDye™ terminators was a major milestone in the evolution of dye terminator sequencing (1,2). Since the two dyes were linked directly together, it was possible to use them for labeling either primers or terminators. The first demonstration that such a doubly-labeled dideoxy terminator could be incorporated by a polymerase was made with a compound called “bifluor” in 1992(3). Improvements in the linkage between the two dyes and changes to the acceptor dyes resulted in improved energy transfer and spectral properties (1,2). Overall, the spare use of the BigDye™ primers and BigDye™ terminators in sequencing reactions resulted in a 4 to 5-fold improvement in signal to noise. The ability to label terminators as well as primers was a significant advantage over previous energy transfer systems for labeling primers that were dependent on an oligonucleotide backbone for energy transfer(4-9).The enzyme that is currently used for sequencing with BigDye™ terminators is AmpliTaq® DNA Polymerase, FS. This enzyme Is a modified form of the DNA polymerase from *Thermus aquaticus* which readily incorporates dideoxy nucleotides and has minimal 5'→3' nuclease activity (10,11,12). Although this enzyme works very well, we have been interested in how other thermostable polymerases would perform in sequencing reactions with the BigDye™ terminators and with a new, more spectrally resolved dye set(13). We have recently begun testing modified DNA polymerases from a *T. thermophilus* strain and from Thermatoga maritima as well as additional variants of Taq DNA Polymerase. The *T. thermophilus* strains differ from the other *Thermus* strains in that the T. *thermophilus* strains are all halotolerant. A distinguishing characteristic of the strain is the ability to grow in medium that contains 3.0% NaCl (15). *Thermatoga maritima* is one of several *Thermotoga* species that have been isolated. The organism is found in marine environments. Unlike the *Thermus* enzymes, the wild-type polymerase from *Thermatoga maritima* has an active 5'→3' as well as 5'→3' nuclease. The enzyme used here has been modified to eliminate this activity as well. Several variants of Taq that show differences in their ability to incorporate BigDye™ terminators have been included in this study.

References

1. Rosenblum, B.B., Lee, L. G., Spurgeon, S. L., et al. 1997. New dye-labeled terminators for improved DNA sequencing patterns. *Nucleic Acids Res*earch 25:4500-4504.
2. Lee, L.G., Spurgeon, S.L., Heiner, C.R., et.al. 1997.New energy transfer dyes for DNA sequencing. *Nucleic Acids Res*earch 25:2816-2822.
3. Lee, L.G., Connell, C.R., Woo, S.L., et.al. 1992. DNA sequencing with dye-labeled terminators and T7 DNA polymerase: effect of dyes and dNTPs on incorporation of dye-terminators and probability analysis of terminatorion fragments. *Nucleic Acids Research* 20:2471-2483.
4. Ju, J., Ruan, C., Fuller, C.W., et.al. 1995. Fluorescence energy transfer dye-labeled primers for DNA sequencing and analysis. *Proc. Natl. Acad. Sci. USA* 92:4347-4351.
5.Ju, J., Kheterpal, I., Scherer, J.R., et.al. 1995. Design and synthesis of fluorescence energy transfer dye-labeled primers and their application for DNA sequencing and analysis. *Analytical Biochemistry* 231:131-140.
6. Hung, S-C., Ju, J., Mathies, R.A., et. al. 1996. Energy transfer primers with 5- or 6-carboxyrhodamine-6G as acceptor chromophores. *Analytical Biochemistry* 238:165-170.
7. Ju, J., Glazer, A.N., Mathies, R.A.1996. Cassette labeling for facile construction of energy transfer fluorescent primers. *Nucleic acids Research* 24:1144-1148.
8. Hung, S-C., Mathies, R.A., Glazer, A.N.1997. Optimization of spectroscopic and electrophoretic properties of energy transfer primers. *Analytical Biochemistry* 252:78-88.
9. Hung, S-C., Mathies, R. A.,Glazer, A. N.1998. Comparison of fluorescence energy transfer primers with different donor-acceptor dye combinations. *Analytical Biochemistry* 255:32-38.
10. Spurgeon, S.L., Reichert, F.L., Koepf, S.M., et.al. 1995. Improved fluorescent dye primer and dye terminator sequencing using a new mutant of Taq DNS Polymerase [abstract]. *Genome Science and Technology* 1:48.
11. Reichert, F.L., Kalman, L.V., Wang, A.M., et.al. 1995. Improved automated DNA cycle sequencing with mutant thermostable DNA polymerases [abstract]. *Genome Science and Technology* 1:46.
12. Kalman, L.V., Abramson, R.D., Gelfand, D.H.1995. Thermostable DNA polymerases with altered discrimination properties [abstract]. *Genome Science and Technology* 1:42.
13. Bond, D.R., Lee, L.G, Rosenblum, B.B., et.al. 2000. New BigDye terminators for automated DNA sequencing [abstract].Presented at 12th International Genome Sequencing and Analysis Conference, Miami Beach Sept. 12-15.
14. Bolchakova, E.V., Spurgeon, S.L., Bone, D.R., et.al. 2000. Comparison of the properties of several bacterial DNA polymerases and their use in fluorescent DNA sequencing [abstract]. Presentedd at 12th International Genome Sequencing and Analysis Conference, Miami Beach, Sept. 12-15.
15. Manaia, C.M.,Hoste, B.,Gutierrez, M. C. and et.al. 1994. Halotolerant thermus strains from marine and terrestrial hot springs belong to *Thermus thermophilus* (ex Oshima and Imahori, 1974) num. rev. emend. *System. Appl. Microbiol*. 17:526-532.

Methods and Materials

Modified DNA polymerases from *Thermus thermophilus* and *Thermatoga maritima* were a gift from David Gelfand at Roche Molecular Systems. Other reagents were from major laboratory suppliers or were supplied by R&D scientists at Applied Biosystems.

Sequencing reactions were done as described for BigDye™ Terminators (ABI PRISM® BigDye™Terminator Cycle Sequencing Ready Reaction Kit Protocol) except that the 8 µL of Ready Reaction Mix was replaced by custom prepared formulations. The cycling program used was as follows: 96°C, 10 seconds, 50°C, five seconds, 60°C for four minutes for 25 cycles unless otherwise noted. Reactions were purified using Centri-Sep™ spin columns from Princeton Separations, Adelphia, NJ. (P/N CS-901).

Data was analyzed on an ABI PRISM® 377 DNA Sequencer using the appropriate run module, matrix and mobility files. The 377 system gels were 36 cm wtr, 36 lanes with 5% Long Ranger® gel unless otherwise noted. Analysis of peak height evenness was done using a custom program developed at Applied Biosystems. Peak height evenness is measured as the relative error: the ratio of the standard deviation to the mean peak height.

Fig. 1 Preliminary testing of new enzymes in sequencing reactions with BigDye™ terminators. The peak pattern obtained with AmpliTaq® DNA Polymerase FS and the modified *Tth* DNA Polymerase are very similar. The peak patterns obtained with modified *Tma* DNA Polymerase and with Taq Variant # 1 show greater differences as compared to the peak patterns obtained with Amplitaq FS. The template in this experiment was pGem® 3Zf(+).

Data obtained with AmpliTaq® DNA Polymerase FS

Data obtained with modified *Tma* DNA Polymerase

Data obtained with modified *Tth* DNA Polymerase

Data obtained with Taq polymerase, variant #1

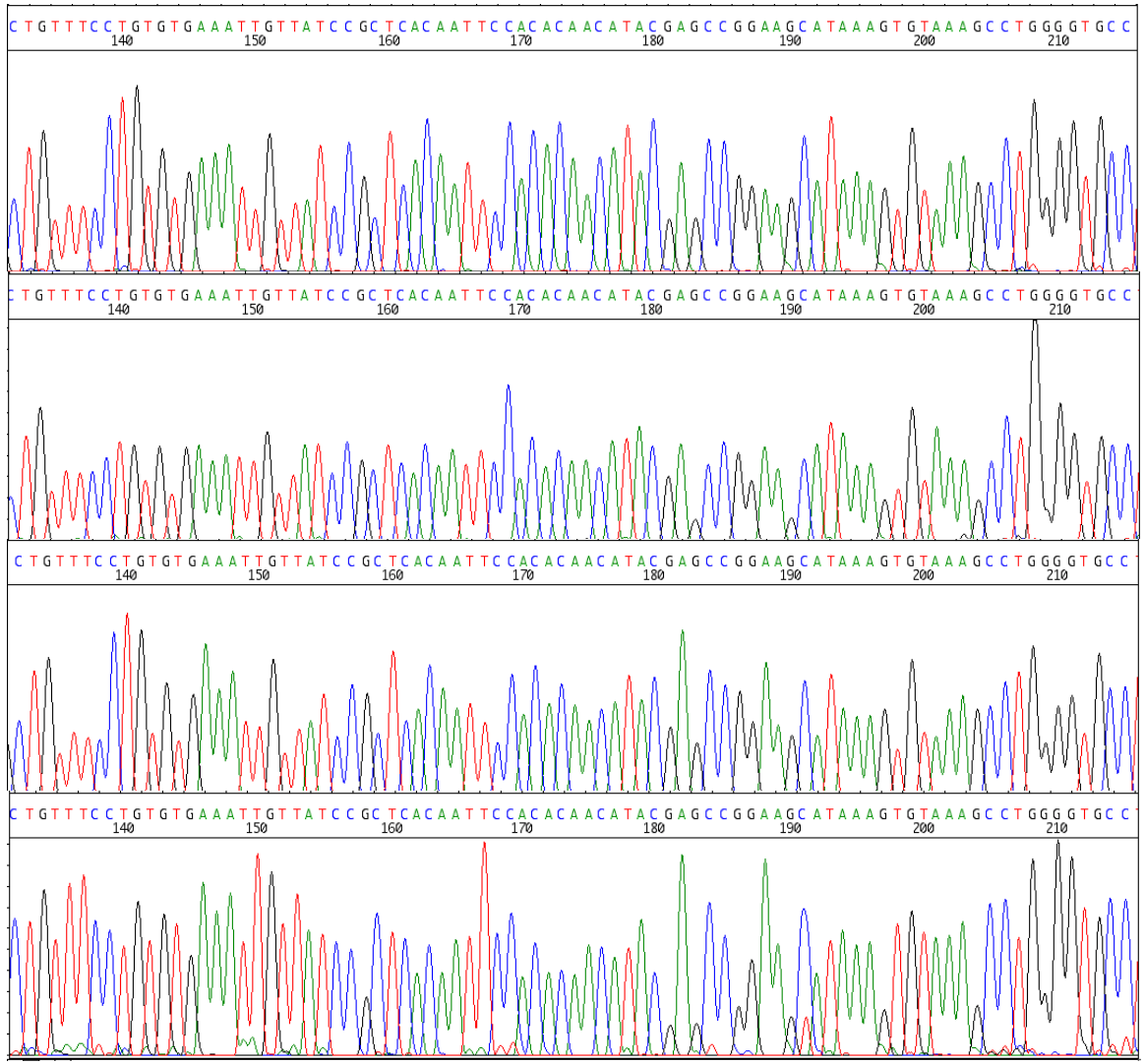


Table 1. Analysis of the sequence data obtained with the various enzymes. Sequence data obtained with pGem3Zf(+) was analyzed for peak height evenness by determining the relative error. Lower numbers indicate more even peak heights. Overall these four enzymes gave similar values. Both the modified *Tma* DNA Polymerase and Variant #1 show a significant improvement in their ability to incorporate pyrimidines relative to purines.

Enzyme	Relative Error				
	G	A	T	C	Average
AmpliTaq® FS	0.355	0.223	0.311	0.308	0.299
Mod. Tma Pol	0.510	0.174	0.287	0.223	0.299
Mod. Tth Pol	0.340	0.320	0.362	0.317	0.335
Taq Variant #1	0.590	0.274	0.322	0.187	0.343

Enzyme	Distribution of Signal				
	%G	% A	%T	% C	Ratio G+A/C+T
AmpliTaq® FS	36.8	17.4	18.7	27.1	1.20
Mod. Tma Pol	15.0	10.99	31.7	42.3	0.35
Mod. Tth Pol	34.6	15.2	23.4	26.8	0.99
Taq Variant #1	18.7	6.9	23.6	50.8	0.34

Table 2. Analysis of data obtained with AmpliTaq® DNA Polymerase,FS, modified *Tma* DNA Polymerase, modified *Tth* DNA Polymerase and Taq Variant #1 with BigDye Terminators. The accuracy of data obtained with modified *Tth* DNA polymerase was very similar to that obtained with AmpliTaq® DNA Polymerase FS and slightly greater than that obtained with modified *Tma* DNA polymerase and Taq Variant #1. The template in these reactions was pGem3Zf(+).

Accuracy

Errors + Ns

Enzyme	Signal	21-120	121-220	221-320	321-420	421-520	521-620	621-720	# Bases with* 99% accuracy
AmpliTaq® FS	3705	0	0	0	0	0	0	3	647
Mod. Tma Pol	3466	0	0	0	0	1	1	3	516
Mod. Tth Pol	5076	0	0	0	0	0	0	4	684
Taq Variant #1	6845	0	0	0	0	0	1	7	605

*Number of bases starting at base 20 that have no more than one error in any 100 base window.

Table 3. Evaluation of other Taq Variants for Sequencing. An additional 7 Taq variants were evaluated for their performance in sequencing. The template used in this experiment was pGem 3Zf(+). The results were compared with those obtained with AmpliTaq® DNA Polymerase, FS. Of the 8 variants tested, only number four failed to extend efficiently. Four of the eight showed improved incorporation of pyrimidines relative to purines. Variants number three, six, seven, and eight gave similar accuracy to that obtained with AmpliTaq® DNA Polymerase FS. Variant number three was selected as giving the best overall performance in this experiment.

Variant	Signal	Purine/Pyrimidine	Accuracy Errors +Ns from 21-620	Ave. Relative Error
#1	3926	0.31	5.6	.362
#2	1212	1.1	6.3	.315
#3	3140	0.38	3.0	.330
#4	139*	----	----	----
#5	3795	0.94	4.3	.377
#6	3341	0.43	2.6	.397
#7	3127	0.40	3.0	.382
#8	2901	0.86	2.6	.429
Amp FS	1948	1.56	2.3	.303

*failed to extend beyond 140 bases.

Fig. 2. Extension in sequencing reactions with GC rich templates with different enzymes. Taq Variants #3 and #1 as well as AmpliTaq® DNA Polymerase FS and modified *Tth* DNA Polymerase were evaluated for their ability to sequence a group of GC rich templates. For pGem3Zf(+) and four of the templates the signal strength was good with all four enzymes. For the most GC rich templates all the enzymes gave sequence data, but with weak signal. Reactions were analyzed on a 377 sequencer with a 48 lane, 36 cm wtr gel.

Enzymes	Lanes
Taq Variant #3	1-8
Taq Variant #1	19-26
Amplitaq FS	27-34
Mod. *Tth* Pol	8-42

Templates	% G+C	Lanes
pGem 3Zf(+)	50.0	1, 19, 27,35
p2358,	62.0	2. 20, 28,36
ABD 70	36.2*	3, 21, 29,37
ABD 3	62.0	4, 22, 30,38
349-1-1	68.9	5, 23, 31,39
4009-1	71.4	6, 24, 32,40
pRY50-8	63.9**	7, 25, 33,41
ABD 32	71.4**	8, 26, 34,42

*contains a region that is 80% G+C
** Based on a partial sequence

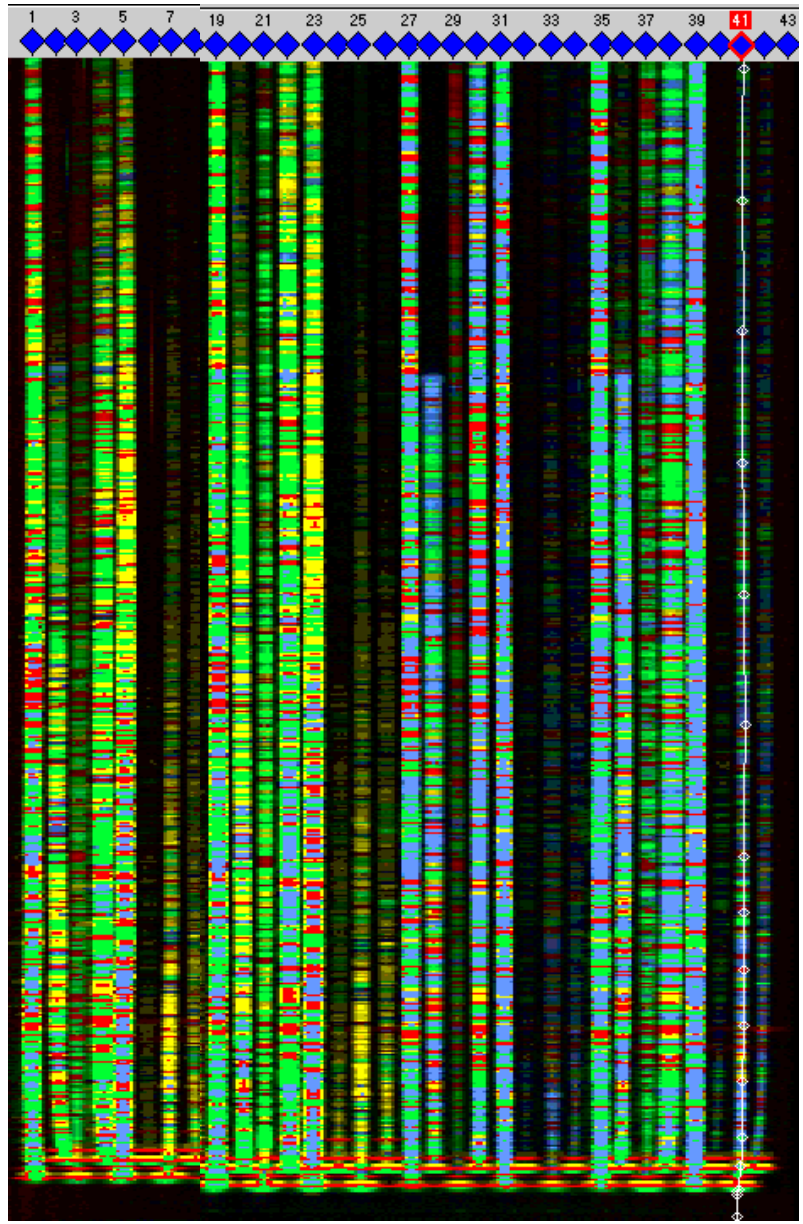
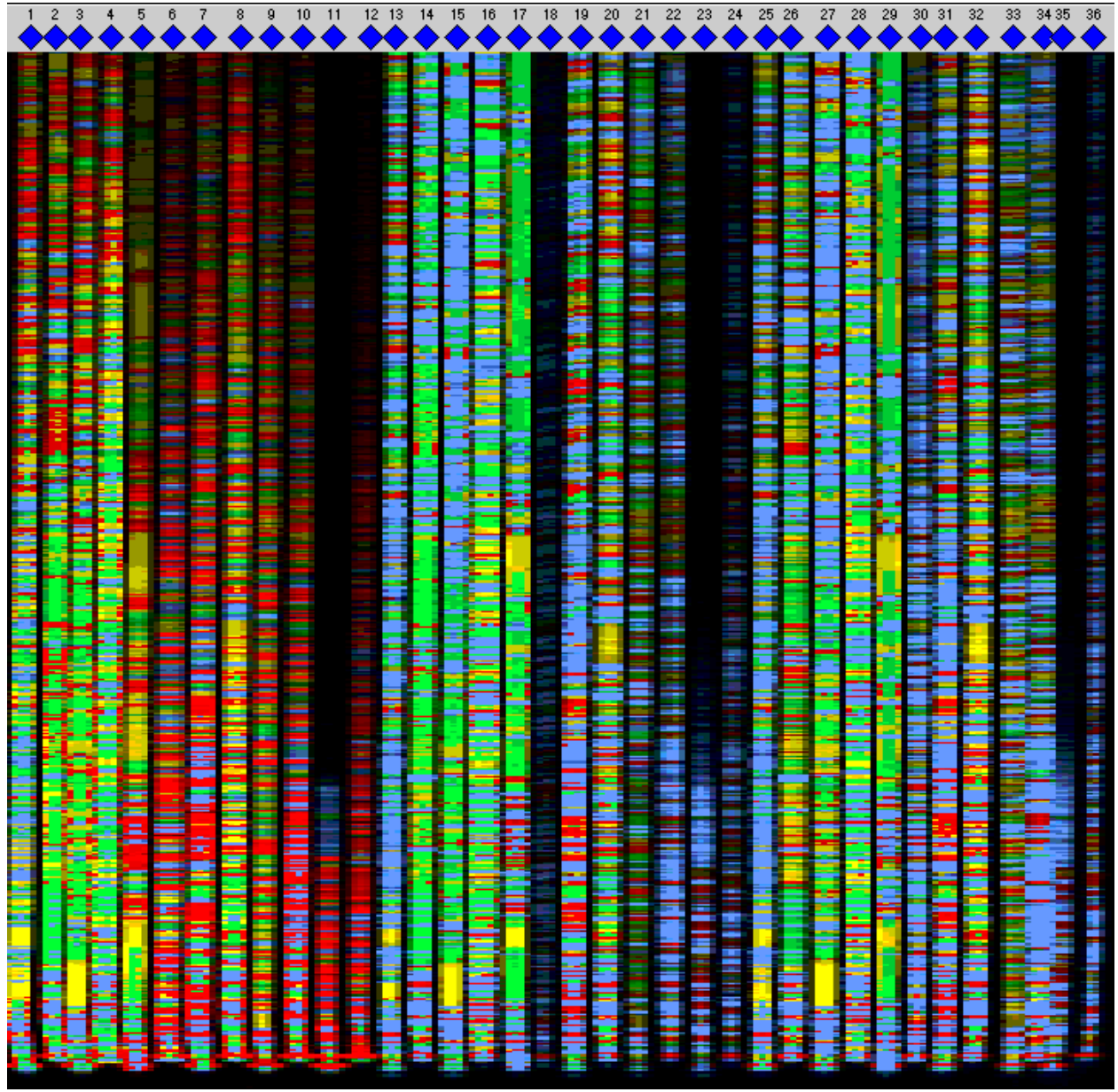


Fig. 3. Sequencing data obtained with a group of AT rich templates. Taq Variant #3, AmpliTaq DNA Polymerase FS and mod. *Tth* DNA Polymerase were also used to sequence a group of AT rich templates. Good data was obtained for all templates except ABD 143 which has a long GA repeat when sequenced in the forward direction. All three enzymes stopped in this region for this template (Lanes 11,23, and 35). Previous work has shown that this template can be successfully sequenced with the dGTP BigDye™ terminator Chemistry.

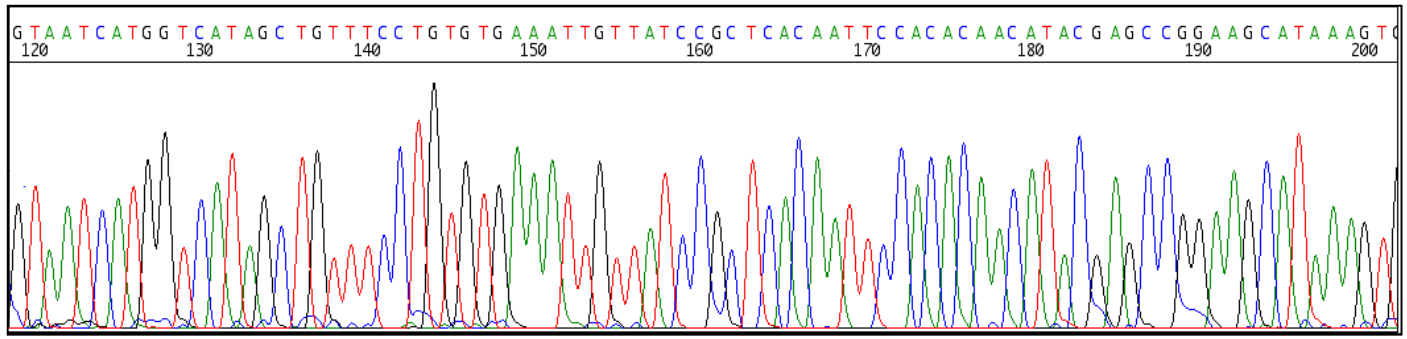


Enzyme	Lanes
Taq Variant #3	1-12
AmpliTaq® DNA Polymerase,FS	13-24
Mod. Tth DNA Polymerase	25-36

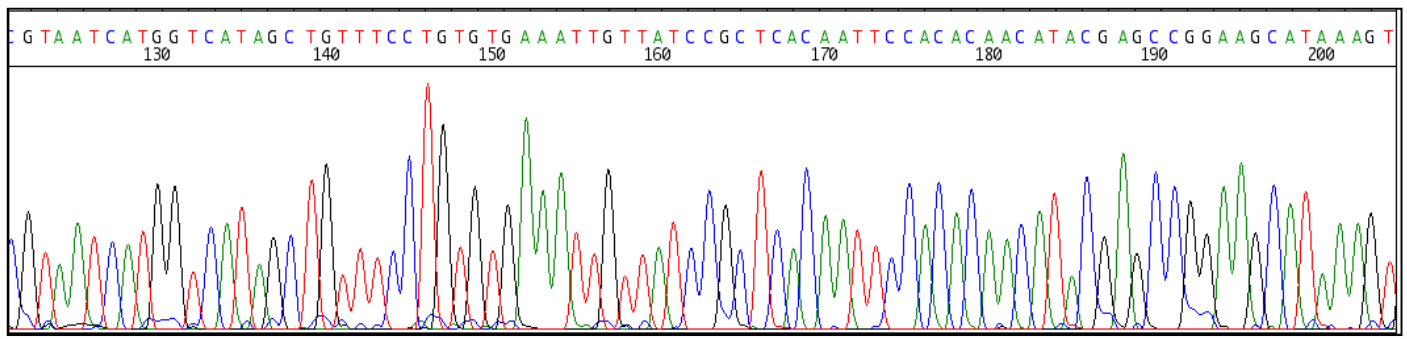
Template	%G+C	Primer	Lanes
DJ2	30.1	-21	1, 13, 25
		Rev	2, 14, 26
pCDNAII1.9	33.0	-21	3, 15, 27
		Rev	4, 16, 28
a-tubulin II	33.3	-21	5, 17, 29
		Rev	6, 18, 30
5B5	36.3	-21	7, 19, 31
		Rev	8, 20. 32
ABD 70	36.2	-21	9, 21, 33
		Rev	10, 22, 34
ABD 143	64.5	-21	11, 23, 35
		Rev	12, 24, 36

Fig. 4. Sequence data obtained with a new, more spectrally resolved dye set. AmpliTaq® DNA Polymerase FS, modified *Tth* DNA polymerase and Taq Variant number three were used in sequencing reactions with the template pGem 3Zf(+). As was observed with the BigDye™ terminators, the peak patterns obtained with AmpliTaq® DNA Polymerase FS and modified *Tth* DNA Polymerase were very similar and noticeably different from that obtained with Taq Variant # 3.

AmpliTaq® DNA Polymerase,FS



Modified *Tth* DNA Polymerase



Taq Variant #3

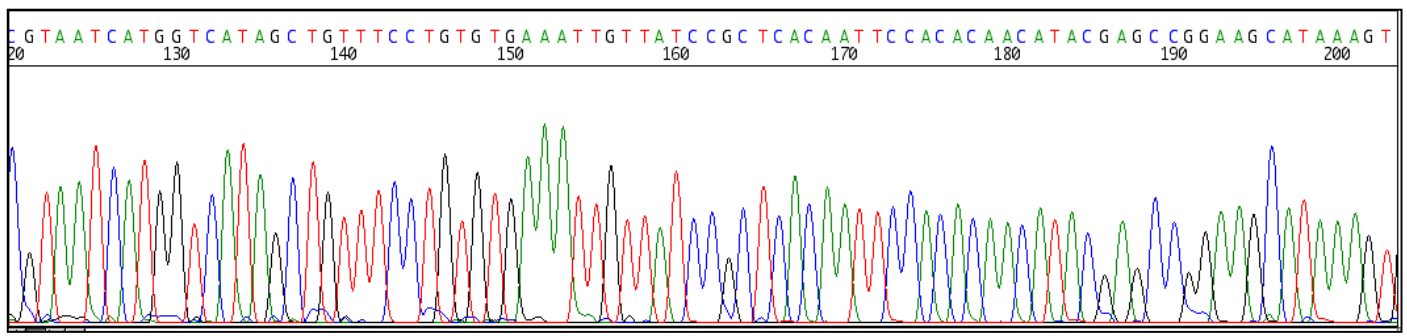


Table. 4. Accuracy of data obtained with a more spectrally resolved dye set. The data shown in Fig. 4 was analyzed for basecalling accuracy. The template in this experiment was pGem® 3Zf(+). All three enzymes gave similar accuracy from base 21-620.

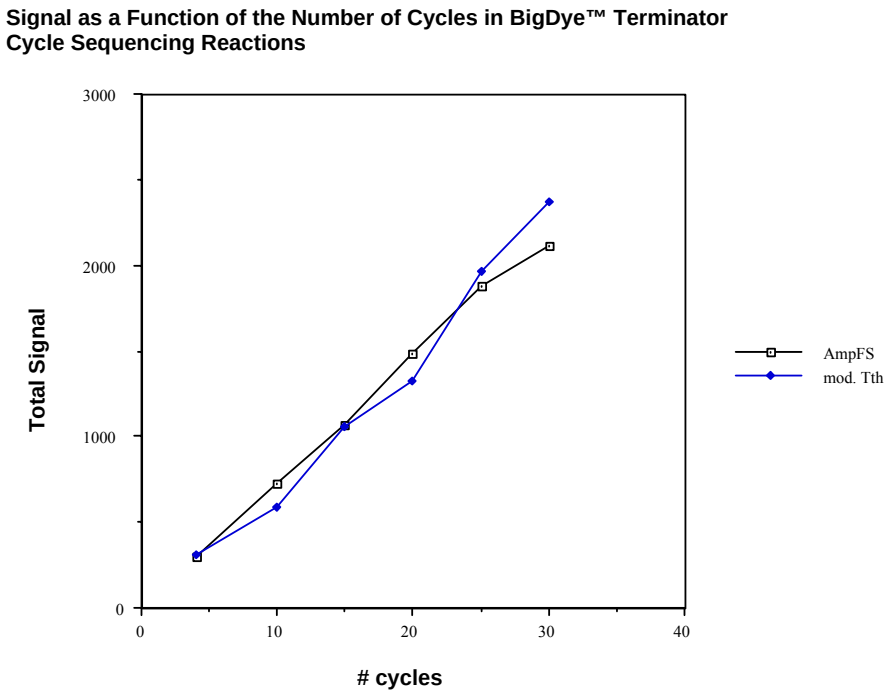
Enzyme	Signal	Errors +Ns							
		1-20	21-120	121-220	221-320	321-420	421-520	521-620	621-720
AmpliTaq® DNA Polymerase, FS	2094	0	0	0	0	0	0	2	11
Mod. *Tth* DNA Polymerase	1740	1	0	0	0	0	0	0	5
Variant #3	2192	1	0	0	0	0	0	0	2

Table 5. Effect of reaction conditions on the signal obtained with AmpliTaq® DNA Polymerase and a modified *Tth* DNA Polymerase. Five templates that differ in their GC content were sequenced with either AmpliTaq® DNA Polymerase, FS or modified *Tth* DNA Polymerase using the standard cycling conditions for BigDye™ terminators, a modified cycling program with 98°C denaturation, or the standard cycling conditions with 5% DMSO included in the reactions. Either modifications increased the signal strength for both enzymes. The success of the approach was dependent on the particular template. For example the use of 98°C denaturation showed little effect with template ABD 77 which has a long CGG repeat, but the signal was significantly improved by the addition of 5% DMSO.

		AmpliTaq® DNA Polymerase, FS			mod. *Tth* DNA Polymerases		
Template	Primer	Standard* Conditions	98**	DMSO***	Standard*	98**	DMSO***
ABD 70	-21	361	633	353	336	1094	583
	Rev	339	761	570	429	1210	N.A.
ABD 78	-21	484	962	1440	587	1466	1876
	Rev	180	1012	1261	249	772	1523
ABD 77	-21	51	51	322	76	82	146
	Rev	86	91	134	139	118	344
4009-1	-21	218	1259	1250	170	1665	1509
	Rev	242	2885	2114	311	3536	3029
ABD 32	-21	167	1695	1397	220	2295	N.A.
	Rev	228	2374	1433	536	3064	2749

•96°C, 10 s, 50°C, 5s, 60°C, 4’
**98 °C, 10 s, 50°C, 5s, 60°C ,4’
***96°C ,10s, 60°C, 5s, 60°C, 4’ with 5% DMSO

Fig. 5. The effect of the number of cycles on BigDye™ terminator reactions. BigDye™ terminator sequencing reactions were carried out for the number of cycles indicated with a 1' extension with either AmpliTaq® DNA Polymerase, FS or modified *Tth* DNA Polymerase. Samples were analyzed on a 377 sequencer with a 36 cm wtr gel. For both enzymes, the signal increased as a function of the number of cycles in an approximately linear fashion up to 30 cycles. The accuracy of the data was similar for both enzymes.



# Cycles	Cumulative Accuracy 21-620 (%)		# Bases with 99% Accuracy	
	Amp FS	**mod*Tth***	**Amp FS**	**mod*Tth***
4	99.5	100.0	575	654
10	99.8	100.0	634	653
15	100.0	100.0	668	688
20	99.7	100.0	698	735
25	99.8	100.0	699	735

Summary

•Modified DNA polymerases from *Thermus thermophilus* and *Thermatoga maritima* have been used successfully in place of AmpliTaq DNA Polymerase, FS in sequencing reactions with BigDye™ terminators.

•Several variants of Taq have been identified that can be used in sequencing reactions with BigDye™ terminators. Several of these show improved incorporation of pyrimidines.

•The enzymes from *T. thermophilus* and one of the Taq variants (variant number three) were also successfully tested with a new set of more spectrally resolved dyes.

•Sequence data obtained with AmpliTaq® DNA Polymerase, FS and modified *Tth* DNA Polymerase were similar in terms of peak pattern evenness and base-calling accuracy. Both enzymes could be used with either a 98° c denaturation or in the presence of 5% DMSO to improve signal with very GC rich templates.

•AmpliTaq® DNA Polymerase, FS and modified *Tth* DNA Polymerase showed a similar increase in signal with increased numbers of cycles up to 30 cycles.

•Further evaluation and optimization of reaction conditions for these enzymes is in progress.

Acknowledgements

The authors would like to thank Dr. David Gelfand from Roche Molecular Systems for providing samples of modified *Tth* DNA Polymerase and modified *Tma* DNA Polymerase used in these experiments.

The authors would also like to thank the organic synthesis group at Applied Biosystems for providing samples of the new dye set used in some of the experiments.