Downloaded from genome.cshlp.org on June 23, 2026 . Published by Cold Spring Harbor Laboratory Press

High-quality automated DNA sequencing primed with hexamer strings.

M A Lodhi and W R McCombie

Genome Res. 1996 6: 10-18
Access the most recent version at doi:10.1101/gr.6.1.10

References This article cites 19 articles, 4 of which can be accessed free at:
http://genome.cshlp.org/content/6/1/10.full.html#ref-list-1

License

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or **click here.**

CRISPR and RNAi Genetic Screening.
Your new superpower.
LEARN MORE
CELLECTA

To subscribe to *Genome Research* go to:
https://genome.cshlp.org/subscriptions

Copyright © Cold Spring Harbor Laboratory Press

Downloaded from genome.cshlp.org on June 23, 2026 . Published by Cold Spring Harbor Laboratory Press

High-quality Automated DNA Sequencing Primed with Hexamer Strings

Muhammad A. Lodhi and W. Richard McCombie[1]

Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724

The finishing phase of genome sequencing projects is expensive, in part, because of the cost of de novo synthesis of custom primers and the management burden associated with obtaining and using them for primer walking. One approach to reduce these high costs is the use of a presynthesized library of short oligonucleotides (8–10 bases) rather than long primers. The use of such a library eliminates the need for custom synthesis of oligonucleotides, providing the convenience of priming from any site by combining two to three short oligonucleotides to form a string with the required specificity. The first practical implementation of this strategy presented a robust protocol for using hexamer strings with radioisotopic labeling. Whereas versions of this technique have subsequently been implemented on fluorescent sequencers we felt that there was a need to develop and extensively test a protocol that consistently gave read lengths comparable to dye-terminator sequencing with longer primers. We have developed a new two-cycle fluorescent Sequenase terminator procedure for using hexamer strings. We tested this procedure using a set of 32 different 3 hexamer primer strings, each known to be functional to some degree in radioisotopic sequencing, on single-stranded M13mp18 template and ABI 373 DNA sequencers. The overall success rate of priming with these hexamer primer strings is 97% with the failure of only one string. In this case, the corresponding 18-mer primer also failed to produce usable sequence from M13mp18 template. The average read length from reactions successfully primed with the 31 different hexamer strings was 461 bases with >99% base-calling accuracy. The current protocol is robust enough to be used in virtually any situation where primer walking on single-stranded templates is used. The success rate and read lengths make it universally applicable to the sequencing of single-stranded templates on automated sequencers. It is also amenable to automation.

In most large-scale genome sequencing projects, a combination strategy that utilizes both random and directed approaches has been employed. In this strategy a random library is made by subcloning cosmid DNA into single- or double-stranded sequencing vectors (Messing et al. 1981; Edwards and Caskey 1991; Andersson and Gibbs 1994, Bodenteich et al. 1994). DNA purified from random subclones is then sequenced using cycle sequencing reactions with fluorescently labeled M13 reverse and universal sequencing primers (Chissoe et al. 1991; Craxton et al. 1991; McCombie et al. 1992). Data generated in this phase of sequencing are assembled, and the remaining gaps as well as regions of insufficient sequence coverage and ambiguities are resolved by primer walking (Strauss et al. 1986; Edwards and Caskey 1991). The degree of finishing required is inversely proportional to the amount of data collected in the random phase. This strategy is termed shotgun sequencing when a large number of fragments are randomly sequenced. The directed or finishing part of such a project is still relatively expensive. The high cost associated with finishing is attributable, in part, to the management burden of de novo custom synthesis of oligonucleotides and the actual cost of the oligonucleotides themselves.

Studier (1989) proposed a strategy in which a library of 8- to 10-base oligonucleotides was used for direct sequencing of genomic DNA. Subsequently other suggestions for using modular primers have been made. Siemieniak and Slighton (1990) proposed the use of a library consisting of nonamer primers, whereas Szybalski (1990) suggested ligating two hexamer primers in situ for sequencing reactions. All of these strategies are based on the concept of utilizing a presynthesized oligonucleotide library eliminating the need for custom primer synthesis for each reaction.

The first demonstration of hexamer primers

[1]**Corresponding author.**
E-MAIL mccombie@cshl.org; FAX (516) 367-8874.

6:10–18 ©1996 by Cold Spring Harbor Laboratory Press ISSN 1054-9803/96 $5.00

Downloaded from genome.cshlp.org on June 23, 2026 . Published by Cold Spring Harbor Laboratory Press

to carry out sequencing reactions was done by Kieleczawa et al. (1992). They used radioactive labeling and single- or double-stranded M13 or M13mp18 template that was saturated with single-strand DNA-binding protein (SSB) from *Escherichia coli*. Subsequently, Kotler et al. (1993) used a combination of pentamers and heptamers, and Azhikina et al. (1993) used strings of pentamers consisting of modified C and A bases to prime sequencing reactions in the absence of SSB, again employing radioisotopic detection.

To make hexamer sequencing technology amenable to automation the above protocols were modified to use fluorescently labeled dye terminators for analysis on automated sequencers (Hou and Smith 1994; Kotler et al. 1994; McCombie and Kieleczawa 1994). Although this work was significant, we felt it necessary to demonstrate adequately the universality of the approach with fluorescent sequencing. We have extensively modified one of these original procedures (McCombie and Kieleczawa 1994). A large percentage (65%–75%) of ~2000 hexamers tested with radioactive sequencing gave readable sequences, although in some cases very weak signals were obtained (W. Studier, pers. comm.). We felt it important to test a number of these primers, including those that gave weak signal in radioactive sequencing to determine the utility of hexamer primers with automated sequencers in a quantitative fashion. Here we report the results of these studies—the testing of 32 different hexamer string primers on single-stranded M13mp18 DNA. These results demonstrate a standardized hexamer string primer sequencing protocol for automated sequencers that typically produce 450–500 bases of high quality sequence on 34-cm gels using ABI 373 DNA sequencers. These results show that this procedure can be used in a wide variety of automated sequencing applications on single-stranded templates.

RESULTS AND DISCUSSION

Multiple Extension Cycles Improve the Results from Hexamer Primed Reactions

Previous protocols of McCombie and Kieleczawa (1994) and Hou and Smith (1994) gave acceptable results using automated sequencers. However, we thought that relatively short read lengths obtained with some of these protocols would be problematic for some sequencing applications. This would be especially true for sequencing human genomic DNA because of the presence of Alu repeats. These repeats of ~300 bases would likely cause serious problems with the modest read lengths obtained previously in some studies with modular primers. More importantly, we felt the need to carry out a comprehensive quantitative analysis of sequence read length and quality obtained with this procedure.

We were concerned that the signal strength we obtained using modular primers was not enough to allow robust reactions with a larger number of primer strings (McCombie and Kieleczawa 1994). Therefore, we elected to improve it by using multiple primer extensions, commonly called cycle sequencing. Current cycle sequencing protocols use as many as 30 extension cycles with *Taq* DNA polymerase. Because the low temperature requirements of hexamer-based reactions prohibit the use of *Taq* DNA polymerase, such large numbers of cycles are clearly impractical. However, early cycle sequencing protocols used far fewer cycles while still improving results over single extension chemistry (Murray 1989). We reasoned that similar results might be obtained with limited cycling of hexamer-primed reactions on automated sequencers. In the protocol we developed a reaction cycle consists of three steps; (1) annealing and initial extension on ice; (2) additional extension and termination at room temperature; and (3) product denaturation at elevated temperature. At the start of each cycle the hexamer string was stabilized on ice and Sequenase polymerase was added. When the reaction is run for one cycle, lower signal results in shorter sequences and more ambiguous base calls (Fig. 1). With the second cycle of amplification, signal intensity is enhanced, without increase in noise, resulting in an improved signal-to-noise ratio (Fig. 1). Sequence length and quality obtained in two-cycle reactions are superior to single extension reactions. Performance of a third cycle improves the signal but does not appreciably increase the read length for the primer strings tested. Because the improvement in quality with a third cycle was marginal overall, we decided to primarily use two cycle reactions and only used three cycles in cases where sequence quality was poor. This minimizes costs and simplifies the reactions. The experiment shown in Figure 1 was performed with a hexamer primer string S06, as well as with a corresponding 18-mer primer, S06L. Because of the similarlity of the effect of number of cycles on the sequence quality, only

Downloaded from genome.cshlp.org on June 23, 2026 . Published by Cold Spring Harbor Laboratory Press

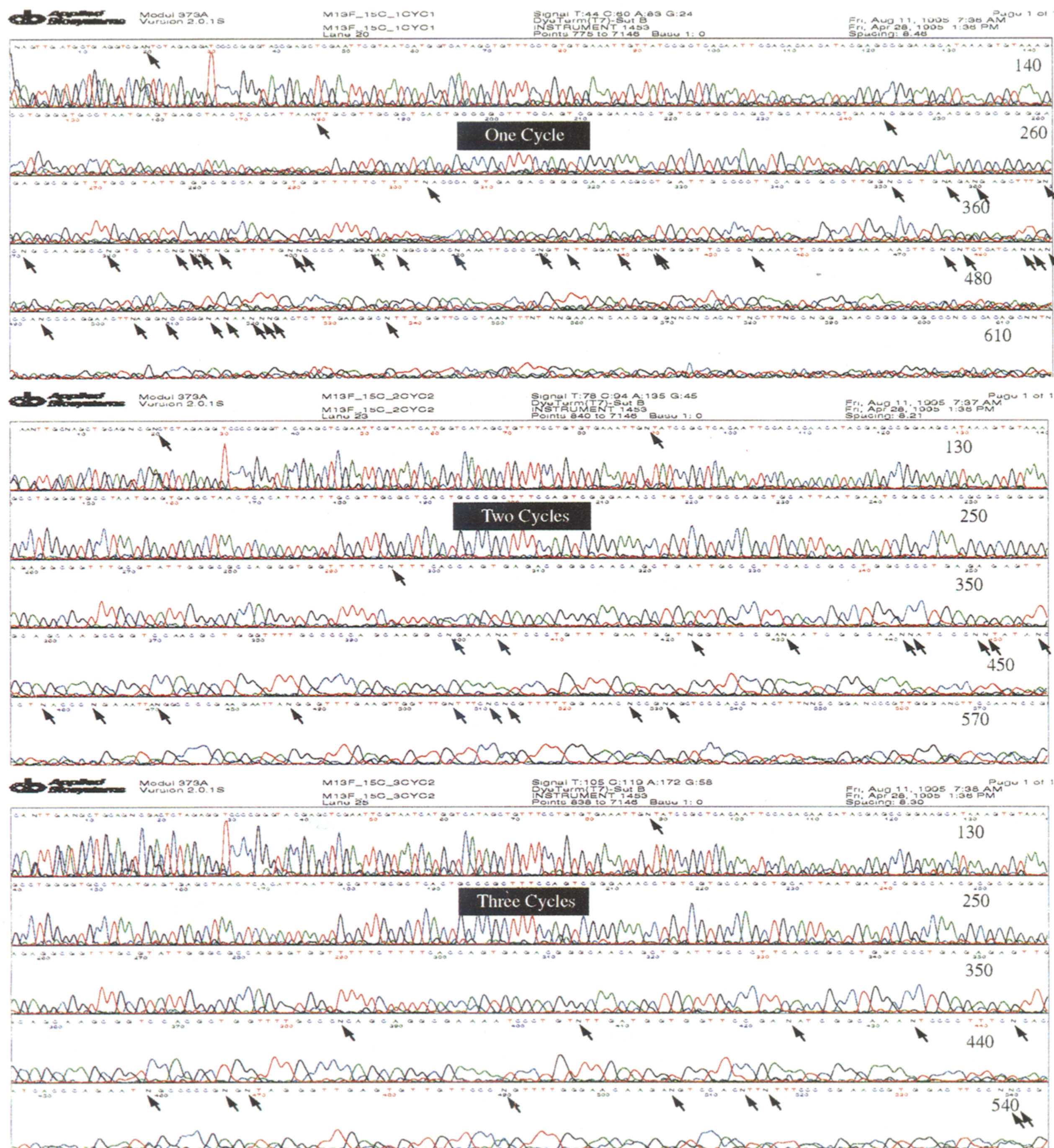

Figure 1 Effect of the number of cycles of Sequenase terminator reaction on sequence read length and quality of the sequence generated by M13 universal primer (5′-TGTAAAACGACGGCCAGT-3′). The reaction was resolved in a 34-cm 4.75% acrylamide gel on an ABI 373 DNA sequencer. The data were collected with ABI collection software (version 1.2.1) using filter set B and analyzed with Analysis software (version 2.0.1) using standard base-calling. The chromatogram at the *top* is the sequence obtained with a single cycle, that in the *middle* with two cycles, and the one at the *bottom* with three cycles of Sequenase terminator reaction. Black arrows indicate ambiguous base calls "N" as detected by Analysis software from base 20–540. Similar experiments of cycle sequencing performed with hexamer string primer S06 resulted in sequence chromatograms of similar quality.

Downloaded from genome.cshlp.org on June 23, 2026 . Published by Cold Spring Harbor Laboratory Press

the chromatograms obtained with S06L have been shown.

It seems likely that the reaction products are denatured from their templates and a second cycle of annealing and extension has occurred. However, a formal possibility exists that the enzyme replenishment alone would enhance sequence quality. We are currently attempting to distinguish between these alternate mechanisms. Based on our initial studies the addition of new enzyme without a heat denaturation step improves read length but the sequence background seems higher and more ambiguities are observed in the sequence. This would indicate that although enzyme replenishment alone improves the reaction, temperature cycling is crucial for optimum results. To test the universality of our current procedure, we elected to try it on a large number of different hexamer strings and compare the sequencing results to those obtained with the corresponding long primers. Moreover, in light of recent improvements in oligonucleotide synthesis, we thought it necessary to compare hexamer read length to the corresponding long primers. The reduced cost of primers makes it important that hexamers deliver comparable results to justify their implementation in the overall sequencing process.

Success Rate and Read Lengths Obtained Using Hexamer Primers on a Single-stranded Template

We tried 32 different hexamer primer strings in Sequenase terminator sequencing reactions with our new protocol. Each string was used multiple times. Of these strings, 31 produced usable sequence. The average read length for each string that was successful was between 315 bases (S30) and 610 bases (S13) (Table 1). Average sequence read length of multiple repetitions of a subset of the corresponding long primers varied between 411 bases (S06L) and 590 bases (S15L). The average read length of the 31 successful strings was 461 bases, whereas read lenghts obtained from reactions primed with eleven 18-base primers with the identical sequence to the corresponding hexamer strings was 492 nucleotides. Part of this difference in the average read lengths of the reactions run with hexamers versus long primers was a result of several short reads obtained from hexamer reactions while conditions were being optimized for amounts of primer, SSB, terminator mix, and Sequenase as well as cycling temperatures in the early part of the project. More recent results with the current version of the procedure yielded consistently equal or higher read lengths when hexamer string primer reactions were compared to similar reactions with longer primers (data not shown).

The success rate obtained when sequencing with the 32 hexamer string primers and Sequenase terminator chemistry was 97%, although primer S01 required a modification of the procedure, as described below. Of all the hexamer strings tested, only S12 and S01 failed to give usable sequence with the standard procedure. Several factors could account for these failures, such as contamination or degradation of the hexamer primers, the existence of secondary structure on the template, or the presence of multiple efficient priming sites. The possibility of contamination or degradation of the hexamers was minimized by obtaining hexamers comprising these two strings from a different commercial source, which did not change the results. To test whether the failure of the priming by the two hexamer strings, S01 and S12, was attributable to the presence of secondary structures, we heated M13mp18 template at 95°C for 2 min and cooled it to ~65°C, and SSB was added followed by running a standard hexamer primer sequencing reaction. Using this method, priming with S01 was successful, resulting in an average read length of 435 bases while S12 still failed. We tested the 18-mer primer (S12L) corresponding to hexamer string S12, and this primer also failed to yield usable sequence. Interestingly, S01 was used successfully, as described in McCombie and Kieleczawa (1994). It is not clear which step in the modified procedure described here led to the difficulties in using this string. Similarly, the reason for the complete failure of S12 and S12L is not understood.

Effects Related to Primer Composition

In the present project hexamer primers were picked from various sites on M13mp18 but mostly from the region between bases 6100 and 6760, where 23 primer strings were picked. Several of these strings differed from each other by only 1 or 2 bases, for example, S21, S24, and S18 or S01, S02, S03, S04, and S05. This strategy of picking primers was adopted to test whether priming could be started at several points in a close region of a genome and whether there is an effect of base composition and stacking at the primer annealing site as reported by Kotler et al.

Downloaded from genome.cshlp.org on June 23, 2026 . Published by Cold Spring Harbor Laboratory Press

Table 1. Primer Sequences, Priming Site on M13 and Sequence Read Length Obtained with Sequencing Reactions Run with Hexamer Primer String on ABI Sequencer Model 373

String name	Primer sequence (5′ → 3′)	M13 site	Read length (range)§	No. of reads	Average read length
Hexamer Primer Strings					
	Hex1 Hex2 Hex3‡				
S01¶	ATACCA/GTCAGG/ACGTTG	787	410-460	2	435
S02	TACCAG/TCAGGA/CGTTGG	786	340-600	5	506
S03	ACCAGT/CAGGAC/GTTGGG	785	360-560	4	475
S04	AGTCAG/GACGTT/GGGAAG	782	230-580	7	447
S05	CAGTCA/GGACGT/TGGGAA	783	230-450	6	353
S06*	TGTAAA/ACGACG/GCCAGT	6291	250-540	9	431
S07	ATCATA/CAGGCA/AGGCAA	7221	390-600	8	460
S08	AGCTCA/TTTTTT/AACCAA	6716	460-600	4	520
S09	AACCGT/GCATCT/GCCAGT	6539	400-600	4	492
S10	CTGTTG/GGAAGG/GCGATC	6406	280-500	3	423
S11	GGGACG/ACGACC/GTATCG	6516	130-510	3	383
S12∫	TGTTGG/GAAGGG/CGATCG	6405	-	-	-
S13	GGCCTC/AGGAAG/ATCGCA	6498	600-620	2	610
S14	TCACAA/TTCCAC/ACAACA	6169	550-620	2	585
S15	GTAAAC/GTTAAT/ATTTTG	6764	300-580	4	447
S16	GCCAGG/GTTTTC/CCAGTC	6316	250-480	3	356
S17	GGCAAA/GCGCCA/TTCGCC	6439	580-580	2	580
S18	CAGCTG/GCGAAA/GGGGGA	6363	520-550	2	535
S19	CCTGCA/GGTCGA/CTCTAG	6259	470-480	2	475
S20	GTGCGG/GCCTCT/TCGCTA	6387	400-520	2	460
S21	GCTGGC/GAAAGG/GGGATG	6361	510-510	2	510
S22	TGCAGG/TCGACT/CTAGAG	6257	540-550	2	545
S23	CCTCTT/CGCTAT/TACGCC	6380	500-520	2	510
S24	AGCTGG/CGAAAG/GGGGAT	6362	510-520	4	512
S25	CAAGGC/GATTAA/GTTGGG	6338	470-480	6	480
S26	GAAGCA/TAAAGT/GTAAAG	6142	280-570	5	444
S27	GTGCCT/AATGAG/TGAGCT	6118	400-490	2	445
S28	ATTTTC/ATCGTA/GGAATC	3418	210-530	4	372
S29	TAATCA/TGGTCA/TAGCTG	6212	370-430	2	400
S30	GGGTGC/CTAATG/AGTGAG	6120	270-390	6	315
S31	TAGAAG/GCTTAT/CCGGTA	3368	210-490	7	408
S32	AGCATA/AAGTGT/AAAGCC	6140	340-430	5	386
Long primers corresponding to hexamer primer strings					
S04L	AGTCAGGACGTTGGGAAG	782	330-600	3	500
S06L	TGTAAAACGACGGCCAGT	6291	400-430	6	411
S07L	ATCATACAGGCAAGGCAA	7221	330-540	4	460
S08L	AGCTCATTTTTTAACCAA	6716	360-560	4	502
S12L∫	TGTTGGGAAGGGCGATCG	6405	-	-	-
S13L	GGCCTCAGGAAGATCGCA	6498	260-600	6	466
S15L	GTAAACGTTAATATTTTG	6764	590-590	2	590
S16L	GCCAGGGTTTTCCCAGTC	6316	470-490	2	480
S29L	TAATCATGGTCATAGCTG	6212	490-530	2	510
S30L	GGGTGCCTAATGAGTGAG	6120	415-470	2	443
S31L	TAGAAGGCTTATCCGGTA	3368	460-490	2	475
S32L	AGCATAAAGTGTAAAGCC	6140	570-580	2	575

Average read length of these readings can be calculated in two ways: (1) The average of each of the 125 individual reads results in an average read length of 445 bases; alternatively, (2) average of individual read lengths, obtained with a primer string, is calculated and these readings of 31 hexamer primer strings give an overall average read length of 461 bases.
*Primer string from −21 M13 universal primer.
‡Hex1 is the 5′, Hex2 is the internal, and Hex3 is the 3′ primer.
§Based on >99% base-calling accuracy.
¶Hexamer string primer reaction was carried out after heating the template at 95°C for 2 min before adding SSB. The reactions failed five times when standard hexamer protocol was used.
∫Sequencing reaction failed seven times with S12 and two times with S12L.

(1993, 1994), who were able to find a relationship between the success of a modular primer combination and the composition and base-stacking interaction between the primers. Individual hexamers used in primer strings varied between 0% to >80% GC content in the present study (Table 1). The GC content of the strings ranged from a low of 22.2%–66.7%. We have not seen any effect either of base composition or base stacking in our experiments. Of the 16 possible base combinations between the 5′ and internal hexamer we successfully tried 14. Similarly, 14 of the possible base combinations between the internal and 3′ hexamer in the string were successfully used. As suggested by Kieleczawa et al. (1992),we avoided using any hexamer string consisting of a hexamer with no GC residues at either the 5′ or 3′ position in the string. We successfully used a hexamer primer consisting only of T's as the internal primer in S08.

Sequencing with Two Hexamers

To determine whether we could simplify these reactions further we tested our

Downloaded from genome.cshlp.org on June 23, 2026 . Published by Cold Spring Harbor Laboratory Press

Table 2. Comparison of Sequence Read Length Generated by Sequenase Terminator Sequencing Reactions Primed with Two and Three Hexamer Primers on M13mp18 Templates

Hexamer string	Hex 1/Hex 2/Hex 3 5′ → 3′	Hexamer primer	Read length (av. of 2)
S02	TACCAG/TCAGGA/CGTTGG	Hex1 and 2	380
		Hex2 and 3	550
		Hex1, 2 and 3	540
S03	ACCAGT/CAGGAC/GTTGGG	Hex1 and 2	Failed
		Hex2 and 3	450
		Hex1, 2 and 3	490
S04	AGTCAG/GACGTT/GGGAAG	Hex1 and 2	450
		Hex2 and 3	480
		Hex 1, 2 and 3	455
S06*	TGTAAA/ACGACG/GCCAGT	Hex1 and 2	465
		Hex2 and 3	495
		Hex1, 2 and 3	480
S07	ATCATA/CAGGCA/AGGCAA	Hex1 and 2	412
		Hex2 and 3	430
		Hex1, 2 and 3	435

Reactions from a string were done at the same time and run on the same gel to minimize variability.
*Primer string from M13 universal primer site.

ability to sequence with two rather than three hexamers. Our attempts to sequence with individual hexamer primers of S06 failed. However, nine strings of two hexamers each were used successfully, and average read lengths varied between 380 and 550 bases. In several cases, this was better than the read lengths obtained with three hexamers (Table 2). Kieleczawa et al. (1992) also observed that two hexamer strings were able to successfully prime radiolabeled reaction. Nine reactions carried out with two hexamer string primers proved successful, whereas one (S03) failed. It appears that in many cases our current protocol is sufficiently robust to carry out sequencing reactions using only two hexamers when analyzed on ABI 373 sequencers. We are continuing our investigation of priming with two hexamers.

Conclusions

The cycling of Sequenase terminator reactions has proved to have a significant advantage over single extension reactions because of enhanced signal strength. Improvement in the signal was noticed particularly in areas where a single cycle produced poor results and thus failed to provide data that could be read accurately by the ABI 373 sequencers. Our initial attempts at developing a cycling protocol simply added an additional reaction cycle to the procedure published previously by McCombie and Kieleczawa (1994). This change alone resulted in greater signal strength and better sequence reads (M.A. Lodhi and W.R. McCombie, unpubl.). The subsequent changes we made to the reaction protocol, like addition of pyrophosphatase, room temperature extension, and reduction of the amount of terminator mix and Sequenase polymerase may also contribute to the improved results we observe with the current protocol. The primary reason for attempting to reduce the amount of terminator mix and Sequenase used, however, was to reduce cost rather than to improve quality. Our overall base-calling accuracy from 31 individual read lengths obtained from hexamer string primers is >99%, with an average read length of 461 bases. We found that a large number of different hexamers could successfully prime reactions using our current protocol. Our current results demonstrate a protocol that provides robust results demonstrating that a number of hexamers had average read lengths sufficient to read through single copies of Alu repeats in human genomic DNA.

At present, one disadvantage of using cycled Sequenase terminator reactions is that enzyme has to be added in each cycle. The availability of a thermostable Sequenase DNA polymerase or a similar enzyme would eliminate this need. If such an enzyme is not available, the automation of the addition of the enzyme would be a straightforward matter. Moreover, we now denature at 65°C rather than at 95°C as in previous protocols. This increases the number of different polymerases that can potentially be used in these reactions, and one of them may be stable at 65°C and active at 4°C. We are currently investigating this possibility.

Downloaded from genome.cshlp.org on June 23, 2026 . Published by Cold Spring Harbor Laboratory Press

Another drawback with the current procedure is the cost of SSB. The cost of commercially obtaining SSB is about $5 per reaction. Although this is considerably less than the cost of an oligonucleotide, it would be valuable to reduce or eliminate this cost. SSB is fairly easy to purify using the protocol of Lohman et al. (1986). In analyzing the mechanism of priming with modular primers, Beskin et al. (1995) found that polymerase preferentially primes using longer primers, modular or not, rather than shorter ones. Based on this observation, SSB may not be required by many primer strings as priming by individual hexamers would be less favored than priming by the complete string. Further work to determine the requirements for SSB in our current protocol (or modified versions of it) is in progress.

At present, a complete hexamer primer library, consisting of 4096 primers, can be purchased for ~$3500. Such a library would provide 10–20 primings with each hexamer. On average, each 18-mer primer in a primer walking reaction costs ~$26. Hence, synthesis of 135, 18-mer primers costs about the same as a full hexamer library. Most of these long primers would not be used more than once in a large-scale sequencing project or in a DNA sequencing core facility, whereas the individual hexamers could be used repeatedly in different combinations. This clearly would result in substantial cost savings for large projects or in core sequencing facilities.

In addition to the cost of oligonucleotides, a significant management burden will be eased by using hexamer strings. Because each oligonucleotide needed is premade, one only needs to combine two or three appropriate hexamers to "synthesize" the desired primer. This makes the custom primers essentially like premade universal and reverse sequencing primers from the standpoint of availability, and offers significant project management and automation advantages over standard primer walking.

These results indicate that our current reaction protocol can be used on a fairly large scale in the end stages of a sequencing project. This step follows when gaps are closed and areas of inadequate coverage or ambiguities are resolved. In most cases, these reactions do not require extremely long read lengths, making our hexamer protocol ideal. Currently in our hands the sensitivity of the ABI sequencers is not high enough to directly sequence high molecular weight templates (e.g., cosmids) using primer walking. We are still pursuing the ultimate goal of direct primer walking on large templates. We feel that the current technology is mature enough to be used in production sequencing for gap filling and conflict resolution. We are currently testing its use in this capacity in the sequencing of three cosmids from fission yeast.

This protocol should facilitate the finishing step of sequencing as well as any other project requiring primer walking on single-stranded templates such as the sequencing of cDNAs. As such, it provides a valuable addition to currently available sequencing tools. Complete automation of these reactions, which our protocol was designed to facilitate, will make them virtually as easy to carry out as universal primer-based reactions.

METHODS

Reagents and Supplies

M13mp18 single-stranded DNA was obtained from New England Biolabs or purified using the protocol of Sambrook et al. (1989). T7 terminator mix containing dNTPαS [ddA–LOU2, 1.2 μM; ddC–5ZOE, 1.8 μM; ddG–5NAN, 3.2 μM; ddT-6FAM 0.44 μM, (Lee et al. 1992)] was purchased from Perkin-Elmer (cat. no. 401489). Manganese chloride was from Sigma (cat. no. M3634). Sequenase version 2.0 T7 DNA polymerase (cat. no. 70775) and polypyrophosphatase (cat. no. 70950) were bought from U.S. Biochemical. MOPS (4 morpholinepropanesulfonic acid) (cat. no. 1124 684) was purchased from Boehringer Mannheim. Proteinase alkaline (Pronase) (cat. no. 539135) was from Calbiochem (cat. no. 539135). SSB protein was manufactured by Pharmacia, Uppsala, Sweden (cat. no. 27-0210-02). Reagent grade ammonium acetate, magnesium chloride, and sodium chloride were from Mallinkrodt. Oligonucleotides were selected from a range of strings shown previously to give a sequencing ladder with radioisotopic sequencing, although in some cases film exposure of several days was required (F.W. Studier and J. Kieleczawa, pers. comm.). Oligonucleotides were synthesized at GIBCO–BRL and the Oligonucleotide Synthesis Facility of Cold Spring Harbor Laboratory. A complete library of 4096 hexamers was purchased from ProtoGene. All of the buffer and chemical solutions were filter sterilized through 0.2-μm Nalgene disposable filter (Nalge Company; cat. no. 120 0020) and stored at –20°C. All sequencing reactions were analyzed on ABI 373 DNA Sequencers (Perkin-Elmer).

Protocol for Sequencing Single-stranded M13mp18 DNA

In a 1.5-ml centrifuge tube, 2 μg of M13mp18 DNA and 7.5 μg of SSB were mixed. Then 3 μl of hexamer primer mix consisting of 1 nmole of each hexamer primer, 2.5 μl of 50 mM $MnCl_2$, 2.5 μl of 150 mM isocitrate in 25% glycerol, and 2.5 μl MOPS buffer [400 mM MOPS (pH 7.5), 500 mM NaCl, 100 mM $MgCl_2$, and 25% glycerol] was added to the

Downloaded from genome.cshlp.org on June 23, 2026 . Published by Cold Spring Harbor Laboratory Press

tube containing the template and SSB. Final volume was adjusted to 20 µl with sterile distilled water and mixed thoroughly. To this mix, 1 unit of Sequenase version 2.0 T7 DNA polymerase containing 0.01 unit of pyrophosphatase and 2 µl of sequenase terminator mixture was added. The reaction was thoroughly mixed and left on ice for 10 min to anneal the primer to the template. Whenever possible following the addition of terminator mix, the reaction tube was protected from light because of the light sensitivity of the terminators. The reaction tube was incubated at room temperature for 15 min for extension and termination of the products. The reaction was then denatured at 65°C for 2 min followed by a 5-min incubation on ice. Following the addition of 1 unit of Sequenase, each tube was incubated on ice for 5 min. The reaction was then transferred to room temperature and incubated for another 10 min for extension and termination. At the end of the second cycle, 2 µg of Pronase was added to each sample. The samples were then vortexed vigorously for 10 sec and heated at 65°C for 10 min to destroy the SSB. The reaction was precipitated with 25 µl of 10 M ammonium acetate and 150 µl 95% ethanol and centrifuged at 12,000 rpm for 20 min at room temperature. The pellet was washed with 300 µl of 70% ethanol, dried briefly in a SpeedVac, resuspended in 3 µl of 5:1 (vol/vol) formamide/50 mM EDTA containing 30 mg/ml of blue dextran, and analyzed on an ABI 373 DNA sequencer at conditions specified by the manufacturer. Alternatively following drying of the pellets, the samples are stored at –20°C until ready for analysis and then resuspended as above.

Sequencing reactions with long primers (18-mers), in place of corresponding hexamer string primers, were run on M13mp18 single-stranded DNA template following the above protocol with the exception that extension/termination was carried out at 37°C instead of at room temperature; 2.5 µl of terminator mix was used and SSB was omitted.

ACKNOWLEDGMENTS

We are thankful to Dr. Jan Kieleczawa (Brookhaven National Laboratory, Brookhaven, NY) for sharing the primer sequences and Dr. Arthur Johnson (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY) for his valuable comments on the manuscript. This project was supported by grant 5R02-HG01045 from the National Institutes of Health.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

References

Andersson, B.R. and A. Gibbs. 1994. In *The polymerase chain reaction* (ed. K.B. Mullis, F.R. Ferre, and R.A. Gibbs), pp. 201–213. Birkhouser, Boston, MA.

Azhikina, T.L., S.V. Veselovskaya, V.A. Myasnikov, V.K. Potapov, O. Ermolayeva, and E.D. Sverdlov. 1993. Strings of contiguous modified pentamers with increased binding affinity can be used for DNA sequencing by primer walking. *Proc. Natl. Acad. Sci.* **90:** 11460–11462.

Bodenteich, S., S. Chissoe, Y.F. Wang, and B.A. Roe. 1994. In *Automated DNA sequencing and analysis* (ed. M.D. Adams, J. Filler, and C. Ventor), pp. 42–50. Academic Press, NY.

Beskin, A.D., D. Zevin-Sonkin, I.A. Sobolev, and L.E. Ulanovsky. 1995. On the mechanism of the modular primer effect. *Nucleic Acids Res.* **23:** 2881–2885.

Chissoe, S.L., W.F. Wang, S.W. Clifton, N. Ma, H.J. Sun, J.S. Lobsinger, S.M. Kenton, J.D. White, and B.A. Roe. 1991. Strategies for rapid and accurate DNA sequencing. *Methods: Companion Methods Enzymol.* **3:** 55–65.

Craxton, M. 1991. Linear amplification sequencing, a powerful method for sequencing DNA. *Methods: Companion Methods Enzymol.* **3:** 20–26.

Edwards, A. and C.T. Caskey. 1991. Closure strategies for random DNA sequencing. *Methods: Companion Methods Enzymol.* **3:** 41–47.

Hou, W. and L.M. Smith. 1994. DNA sequencing with a hexamer string primer and dye-labeled terminators. *Anal. Biochem.* **221:** 136–141.

Kieleczawa, J., J.J. Dunn, and F.W. Studier. 1992. DNA sequencing by primer walking with strings of contiguous hexamers. *Science* **258:** 1787–1791.

Kotler, L., I. Sobolev, and L. Ulanovsky. 1994. DNA sequencing: Modular primers for automated walking. *BioTechniques* **17:** 554–558.

Kotler, L.E., D. Zevin-Sonkin, I. Sobolev, and A.D. Beskin. 1993. DNA sequencing: Modular primers assembled from a library of hexamers or pentamers. *Proc. Natl. Acad. Sci.* **90:** 4241–4245.

Lee, L.G., C.R. Connell, S.L. Woo, R.D. Cheng, B.F. McArdle, C.W. Fuller, N.D. Halloran, and R.K. Wilson. 1992. DNA sequencing with dye-labeled terminators and T7 DNA polymerase: Effects of dyes and dNTPs on incorporation of dye-terminators and possibility analysis of termination fragments. *Nucleic Acids Res.* **20:** 2471–2483.

Lohman, T.M., J.M. Green, and R.S. Beyer. 1986. Large-scale overproduction and rapid purification of the *Escherichia coli ssb* gene product. Expression of *ssb* gene under IP_L control. *Biochemistry* **25:** 21–25.

McCombie, W.R., C. Heiner, J.M. Kelley, M.G. Fitzgerald, and J.D. Gocayne. 1992. Rapid and reliable fluorescent cycle sequencing of double-stranded templates. *DNA Sequencing* **2:** 289–296.

McCombie, W.R. and J. Kieleczawa. 1994. Automated DNA sequencing using 4-color fluorescent detection of reactions primed with hexamer strings. *BioTechniques* **17:** 574–579.

Messing, J.P., H. Crea, and P.H. Seeburg. 1981. A system

Downloaded from genome.cshlp.org on June 23, 2026 . Published by Cold Spring Harbor Laboratory Press

for shotgun DNA sequencing. *Nucleic Acids Res.* **9:** 309–321.

Murray, V. 1989. Improved double-stranded DNA sequencing using the linear polymerase chain reaction. *Nucleic Acids Res.* **17:** 88–89.

Sambrook, J., E.F. Fritsch, and T. Maniatis. 1989. *Molecular cloning: A laboratory manual* 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Siemieniak, D.R. and J.L. Slightom. 1990. A library of 3342 useful nonamer primers for genome sequencing. *Gene* **96:** 121–124.

Strauss, E.C., J.A. Kobori, G. Siu, and L.E. Hood. 1986. Specific-primer-directed DNA sequencing. *Anal. Biochem.* **154:** 353–360.

Studier, F.W. 1989. A strategy for high-volume sequencing of cosmid DNAs: Random and directed priming with a library of oligonucleotides. *Proc. Natl. Acad. Sci.* **86:** 6917–6921.

Szybalski, W. 1990. Proposal for sequencing DNA using ligation of hexamers to generate sequential elongation primers (SPEL-6). *Gene* **90:** 177–178.

Received September 18, 1995; accepted in revised form December 24, 1995.