Benchmarks

Fiber DNA measurements were performed using two different microscope settings (Leica and Olympus) in two different laboratories with identical results. Our data indicate that this fast and simple procedure allows DNA size measurements that are as accurate or more accurate (5%–10%) than those taken by gel electrophoresis.

REFERENCES

1. Allemand, J.F., D. Bensimon, L. Jullien, A. Bensimon, and V. Croquette. 1997. pH-dependent specific binding and combing of DNA. *Biophys. J.* 73:2064–2070.
2. Bensimon, A., A. Simon, A. Chiffaudel, V. Croquette, F. Heslot, and D. Bensimon. 1994. Alignment and sensitive detection of DNA by a moving interface. *Science* 265:2096–2098.
3. Horelli-Kuitunen, N., J. Aaltonen, M.L. Yaspo, M. Eeva, M. Wessman, L. Peltonen, and A. Palotie. 1999. Mapping ESTs by fiber-FISH. *Genome Res.* 9:62–71.
4. Hu, X., C. Aston, and D.C. Schwartz. 1999. Optical mapping of DNA polymerase I action and products. *Biochem. Biophys. Res. Commun.* 254:466–473.
5. Jing, J., Z. Lai, C. Aston, J. Lin, D.J. Carucci, M.J. Gardner, B. Mishra, T.S. Anantharaman et al. 1999. Optical mapping of *Plasmodium falciparum* chromosome 2. *Genome Res.* 9:175–181.
6. Jing, J., J. Reed, J. Huang, X. Hu, V. Clarke, J. Edington, D. Housman, T.S. Anantharaman et al. 1998. Automated high resolution optical mapping using arrayed, fluid-fixed DNA molecules. *Proc. Natl. Acad. Sci. USA* 95:8046–8051.
7. Laan, M., O.P. Kallioniemi, E. Hellsten, K. Alitalo, L. Peltonen, and A. Palotie. 1995. Mechanically stretched chromosomes as targets for high-resolution FISH mapping. *Genome Res.* 5:13–20.
8. Michalet, X., R. Ekong, F. Fougereousse, S. Rousseaux, C. Schurra, N. Hornigold, M. van Slegtenhorst, J. Wolfe et al. 1997. Dynamic molecular combing: stretching the whole human genome for high-resolution studies. *Science* 277:1518–1523.
9. Raap, A.K., R.J. Florijn, L.A.J. Blonden, J. Wiegant, J.W. Vaandrager, H. Vrolijk, J. den Dunnen, H.J. Tanke et al. 1996. Fiber FISH as a DNA mapping tool. *Methods* 9:67–73.
10. Raap, A.K. and J. Wiegant. 1994. Use of DNA-halo preparations for high-resolution DNA in situ hybridization. *Methods Mol. Biol.* 33:123–130.
11. Samad, A.H., W.W. Cai, X. Hu, B. Irvin, J. Jing, J. Reed, X. Meng, J. Huang et al. 1995. Mapping the genome one molecule at a time—optical mapping. *Nature* 378:516–517.
12. Weier, H.U., M. Wang, J.C. Mullikin, Y. Zhu, J.F. Cheng, K.M. Greulich, A. Bensimon, and J.W. Gray. 1995. Quantitative DNA fiber mapping. *Hum. Mol. Genet.* 4:1903–1910.
13. Yokota, H., F. Johnson, H. Lu, R.M. Robinson, A.M. Belu, M.D. Garrison, B.D. Ratner, B.J. Trask et al. 1997. A new method for straightening DNA molecules for optical restriction mapping. *Nucleic Acids Res.* 25:1064–1070.

We thank Drs. Shuming Nie (Indiana University) and William A. Lyon (Affymetrix) for their technological insights regarding the DNA fiber preparation technique. Address correspondence to Dr. Octavian Henegariu, Microarray Facility, MB & B Department, Yale University School of Medicine, WWW 5, 333 Cedar Street, New Haven, CT 06510, USA. e-mail: octavian.henegariu@yale.edu

Received 21 December 2000; accepted 25 April 2001.

O. Henegariu, L. Grober, W. Haskins¹, P.N. Bowers, M.W. State, N. Ohmido², P. Bray-Ward, and D.C. Ward
*Yale University
 New Haven, CT
¹University of Florida
 Gainesville, FL, USA
²Hokuriku National
 Agricultural Experiment
 Station
 Joetsu, Japan*

Method to Introduce Stable, Expanded, Polyglutamine-Encoding CAG/CAA Trinucleotide Repeats into CAG Repeat-Containing Genes

BioTechniques 31:250–254 (August 2001)

Expansion of CAG trinucleotide repeats coding for polyglutamine causes eight progressive neurodegenerative diseases, including Huntington's disease (HD) and a number of inherited spinocerebellar ataxias (SCA) (1). While these disorders are only slowly progressive, investigation of the effects of elongated polyglutamine in cell lines and

transgenic animals can be accelerated by using constructs containing long CAG repeats expanded even beyond the pathological size. Often, the expanded repeats are obtained from patient's material; however, native CAG tracts are unstable in bacteria (4), as well as in stably transfected cell lines (3) and transgenic mice (8). PCR-based methods were developed to enlarge CAG repeats within cDNAs (9), some also introducing CAA interruptions (6), which increase the stability of CAG repeats, while still coding for glutamine. However, misincorporation errors can occur during PCR, and the final repeat size is not always predictable. Recently, Kazantsev et al. (5) described the production of alternating CAG/CAA repeats, (5'-CAACAGCAGCAACAGCAA-3')_n, ranging in size from 25 to 300 triplets, which remained stable in bacteria and in several different eukaryotic systems (unpublished data). Here, we describe a straightforward method to introduce these repeats (20, 41, 60, 67, 98, 185, or 264 units) into the native CAG repeat of any gene. Our strategy is advantageous because the resulting repeats are stable, have a predictable size, and the amino acid sequence surrounding the polyglutamine tract is preserved. While our method can be used for any gene, we elongated the SCA7 CAG repeat to encode 107 and 273 glutamines.

The strategy is outlined in Figure 1. Four bases—GTAC—were introduced into the SCA7 CAG repeat by in vitro mutagenesis to create a *KpnI* site (GGTACC) (Figure 1, left). Since the SCA7 cDNA already contains a *KpnI* site, only a part of the gene [bases 1–654 of the SCA7 open reading frame (ORF) with 9 CAG] was *EcoRI* cloned into pBluescript® SK II(+) vector (Stratagene, La Jolla, CA, USA), in which we previously removed the *KpnI* site from the polylinker (pBSSKII/*KpnI*-). The in vitro mutagenesis was performed using primers SCA7CAG *KpnI* [5'-CGG(CAG)₄GTAC(CAG)₅CCG-3'] and SCA7CTG*KpnI* [5'-CGG(CTG)₅GTAC(CTG)₄CCG-3'], as follows. The SCA7 segments upstream and downstream of the CAG repeat were amplified separately, using SCA7CTG*KpnI* and SCA7CAG*KpnI* primers in combination with vector primers, M13 Reverse and Forward, re-

Benchmarks

spectively. The PCR products, overlapping through the modified CAG repeat, were combined and used as template in a PCR with M13 Forward and Reverse primers. All PCRs were performed using *Pfu* DNA polymerase (Promega, Madison, WI, USA), and 1.5 to 2 M betaine was supplemented to facilitate amplification of GC-rich sequences surrounding the *SCA7* CAG tract (2). The resulting product was *Eco*RI digested to remove the flanking polylinker sequences, cloned back into pBSSKII/*Kpn*I-, and the presence of the *GTAC* insertion was verified by sequencing. Subsequent *Kpn*I digestion, dephosphorylation, and removal of 3' overhangs by Klenow fragment or T4 DNA polymerase created an *SCA7* "acceptor" construct for the expanded CAG/CAA repeats (Figure 1, left).

The CAG/CAA repeats were obtained from EGFP-polyglutamine fusion constructs (5) as *Hind*III-*Pst*I (25, 46, 65, 72, and 103 units) and *Hind*III-*Xho*I (191 and 270 units) fragments and cloned into the respective sites of two different *Gsu*I-linkers, *HP-Gsu*I (5'-GGATCACTGGAGAAGCTTAGACTGGAATCTGCAGTCTTCTCCAGCTAGA-3') and *HX-Gsu*I (5'-GGATCAC-TGGAGAAGCTTAGACTGGAATCTCGAGTCTCCAGCTAGA-3'), respectively, which were inserted between the *Bam*HI and *Xba*I sites of pGEM7zf(+) (Promega). Figure 1 (right) details the *HP-Gsu*I linker. A 10-bp spacer separates the *Hind*III and *Pst*I sites, allowing for efficient double digestion. Furthermore, the CAG/CAA tracts cloned into this linker are flanked by two sites for *Gsu*I, a restriction endonuclease cutting a fixed number of nucleotides downstream of its recognition sequence, located such as to release CAG/CAA repeats, which, after removing their 3' overhangs, can be ligated into the *SCA7* acceptor construct while maintaining the correct reading frame. Importantly, these repeats are several units shorter than the original CAG/CAA tracts cloned in the linkers (i.e., 20, 41, 60, 67, 98, 185, and 264 units long). While we generated *Gsu*I-linker constructs for all of the sizes, we only inserted repeats of 98 and 264 units into the *SCA7* CAG repeat, which thus encoded in total 107 and 273 glutamines. One clone with the correct orientation of the inserted CAG/CAA

Figure 1. Strategy used to introduce expanded CAG/CAA tracts into an existing (SCA7) CAG repeat. (Left) Portion of the *SCA7* cDNA containing the CAG repeat was inserted into pBSSKII(+) vector, in which the *Kpn*I site was removed from the polylinker (pBSSKII/*Kpn*I-). *In vitro* mutagenesis using the indicated primers introduced four bases—*GTAC*—creating a *Kpn*I site within the CAG repeat. *Kpn*I digestion, dephosphorylation, and removal of 3' overhangs generated the *SCA7* acceptor construct. (Right) CAG/CAA repeats as *Hind*III-*Pst*I (25, 46, 65, 72, and 103 units) or *Hind*III-*Xho*I (191 and 270 units) fragments were inserted into the respective sites of specially designed *Gsu*I linkers cloned between the *Bam*HI and *Xba*I sites of pGEM7zf(+) vector. Only the linker containing *Hind*III and *Pst*I (*HP-Gsu*I linker) is shown here. The *Hind*III and *Pst*I sites are separated by a 10-bp spacer to allow for efficient double digestion. Flanking the repeats are two sites for *Gsu*I, positioned such as to release CAG/CAA tracts (20, 41, 60, 67, 98, 185, and 264), which, after removal of their 3' overhangs, can be ligated into the CAG repeat of the *SCA7* acceptor construct.

Benchmarks

Figure 2. SCA7 CAG repeat containing the (CAG/CAA)₉₈ insertion. CAG/CAA repeat of 98 units was inserted into the SCA7 CAG repeat. Only the 5' (A) and 3' (B) sequences flanking the insertion site (arrows) are shown to illustrate the polyglutamine tract encoded by the repeat. Moreover, the (5'-CAACAGCAGCAACAGCAA-3') repetitive consensus of the inserted repeat is underlined in (A).

tract and intact cloning boundaries was identified for each repeat size by sequencing (Figure 2), thus proving the usefulness of our method.

In conclusion, we described a cloning method that allows introduction of different-sized, stable CAG/CAA repeats into an existing CAG repeat of any gene while preserving the amino acid sequence surrounding that repeat. Importantly, this method is fast, its outcome is predictable, and it avoids PCR amplification of the CAG repeat. Our method includes several strategies previously employed by others. For example, digestion with a shifted-site restriction enzyme to release a CAG repeat suitable for further in-frame cloning was used by Peters and Ross (9) in a PCR-based method for expanding existing CAG repeats. Moreover, in recent work, Lin et al. (7) introduced a *Ban*HI site into the CAG repeat of the mouse huntingtin gene, which—after digestion and removal of 5' overhangs—allowed for in-frame insertion of PCR-synthesized CAG repeats. While this latter approach is similar to ours, we preferred using *Kpn*I because the 3' overhangs created by *Kpn*I can be removed easily by T4 DNA polymerase or Klenow fragment, while *Ban*HI-generated 5' overhangs need to be removed by a single-strand specific nuclease (e.g., S1 or mung bean nuclease), which might cause unwanted DNA degradation. Importantly, several other enzymes can be used instead of *Kpn*I (GGTAC'C), such as *Apa*I (GGGCC'C), *Bbe*I (GGCGC'C), *Sph*I (GCATG'C) and *Sac*I (GAGCT'C), all generating 3' overhangs flanked by G and

C. This provides additional flexibility to our method. Clearly, our method combined and advanced previous techniques, allowing scientists studying CAG triplet repeat diseases to rapidly insert characterized and stable CAG/CAA repeats into their cDNAs of interest.

REFERENCES

1. Cummings, C.J. and H.Y. Zoghbi. 2000. Fourteen and counting: unraveling trinucleotide repeat diseases. *Hum. Mol. Genet.* 9:909-916.
2. Henke, W., K. Herdel, K. Jung, D. Schnorr, and S.A. Loening. 1997. Betaine improves the PCR amplification of GC-rich DNA sequences. *Nucleic Acids Res.* 25:3957-3958.
3. Jana, N.R., M. Tanaka, G. Wang, and N. Nukina. 2000. Polyglutamine length-dependent interaction of hsp40 and hsp70 family chaperones with truncated N-terminal huntingtin: their role in suppression of aggregation and cellular toxicity. *Hum. Mol. Genet.* 9:2009-2018.
4. Kang, S., A. Jaworski, K. Ohshima, and R.D. Wells. 1995. Expansion and deletion of CTG repeats from human disease genes are determined by the direction of replication in *E. coli*. *Nat. Genet.* 10:213-218.
5. Kazantsev, A., E. Preisinger, A. Dranovsky, D. Goldgaber, and D. Housman. 1999. Insoluble detergent-resistant aggregates form between pathological and nonpathological lengths of polyglutamine in mammalian cells. *Proc. Natl. Acad. Sci. USA* 96:11404-11409.
6. Laccone, F., R. Maiwald, and S. Bingemann. 1999. A fast polymerase chain reaction-mediated strategy for introducing repeat expansions into CAG-repeat containing genes. *Hum. Mutat.* 13:497-502.
7. Lin, C.H., S. Tallaksen-Greene, W.M. Chien, J.A. Cearley, W.S. Jackson, A.B. Crouse, S. Ren, X.J. Li, R.L. Albin, and P.J. Detloff. 2001. Neurological abnormalities in a knock-in mouse model of Huntington's disease. *Hum. Mol. Genet.* 10:137-144.

8. Mangiarini, L., K. Sathasivam, A. Mahal, R. Mott, M. Seller, and G.P. Bates. 1997. Instability of highly expanded CAG repeats in mice transgenic for the Huntington's disease mutation. *Nat. Genet.* 15:197-200.

9. Peters, M.F. and C.A. Ross. 1999. Preparation of human cDNAs encoding expanded polyglutamine repeats. *Neurosci. Lett.* 275:129-132.

Address correspondence to Prof. Dr. Christine Van Broeckhoven, Laboratory of Molecular Genetics, University of Antwerp (UIA), Department of Biochemistry, Universiteitsplein 1, B-2610 Antwerpen, Belgium. e-mail: cvbroeck@uia.ua.ac.be

Received 22 January 2001; accepted 11 May 2001.

Andrej Michalik, Aleksey Kazantsev¹, and Christine Van Broeckhoven
*University of Antwerp
Antwerpen, Belgium*
¹*Massachusetts Institute of
Technology
Cambridge, MA, USA*