Vol. 205, No. 1, 1994 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
November 30, 1994 Pages 944-947

CHIMERISM IN PCR PRODUCTS FROM A MULTIGENE FAMILY

D. Bhavsar, Hd. Zheng and J. Drysdale

Department of Biochemistry, Tufts Medical School,
136 Harrison Avenue, Boston, MA 02111

Received October 24, 1994

Abstract. This paper describes a potential hazard in obtaining sequences of a multigene family by amplifying the mixed population by the polymerase chain reaction with primers from conserved sequences. We describe a product which appears to be a chimera formed during the PCR reaction from two different but related sequences. © 1994 Academic Press, Inc.

Amplification of genomic sequences by the polymerase chain reaction (PCR) is a convenient method for obtaining sequences from multigene families without cloning individual members. We applied this approach to obtain sequences from members of the human ferritin H gene family. This family contains about 17 highly homologous members, many of which appear to be processed pseudogenes (1-3). In a search for other functional sequences we amplified H-like sequences from genomic DNA by PCR using a primer set from the known functional gene and cloned and sequenced the products. This paper describes a potential source of error with this approach through the formation of a chimera from two different ferritin genes during the PCR reaction.

Material and Methods

PCR Amplification. Human DNA was obtained from lymphocytes by standard methods. Ferritin H-like sequences were amplified from this DNA with a primer set, 5' TCTTACTACTTTGACCGCGAT 3' and 5' CAACATGCATGCACTGCCTT 3' corresponding to positions 321-341 and 843-822 of the positive and negative strands respectively of the liver H cDNA, pHF16 (1). Reactions contained 1 ug DNA, 0.2uM primer, 1mM dNTP, 5 mM Mg^{2+}. Denaturation was at 94° for 1 min, annealing at 55° for 1 min and extension at 72° for 2 mins. Thirty cycles were performed in a Perkin Elmer Cetus DNA Thermal Cycler.
Sequencing. The PCR products were purified by chromatography on Centricon columns and cloned into the HincII site of Bluescript plasmid. Sequencing was performed on both strands using T3 and T7 primers.

0006-291X/94 $5.00
Copyright © 1994 by Academic Press, Inc.
All rights of reproduction in any form reserved.

Results

We obtained many different H-like sequences from the library of PCR products amplified from genomic DNA with the H primer set. Several were identical to conventionally cloned sequences, confirming the essential validity of the PCR approach. We also obtained several new

```
      330
FTH1   TCT TAC TAC TTT GAC CGC GAT GAT GTG GCT TTG AAG AAC TTT GCC
FTHL1  --- --- -G- --- --- --T --- --- --- --- --- --- --- --- ---
PCR5                               --- --- --- --- --- --- --- ---
FTHP2  --- --- --- --- --- --  --- --- --A --- -A- --T --- --- ---
     375
FTH1   AAA TAC TTT CTT CAC CAA TCT CAT GAG GAG AGG GAA CAT GCT GAG
FTHL1  --- --- --- --- --- --- --- --- --- --- --- --G --- --- ---
PCR5   --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
FTHP2  --- --- --- --- --- --- --- --- --- --- --- -C- --- --C A--
     420
FTH1   AAA CTG ATG AAG CTG CAG AAC CAA CGA GGT GGC CGA ATC TTC CTT
FTHL1  --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
PCR5   --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
FTHP2  GG-A--- --- --- --- --- --- --- --- --- --- -X- --- --- ---
     465
FTH1    CAG GAT ATC AAG AAA CCA GAC TGT GAT GAC TGG GAG AGC GGG CTG
FTHL1   --- --- --- --A --- --- --- --- --- --- --- --- --- --- ---
PCR5    --- --- --- --A --- --- --- --- --- --- --- --- --- --- ---
FTHP2   --A --- --- --- --- --- --- --- --- --- --- --- --- T-- ---
     510
FTH1   AAT GCA ATG GAG TGT GCA TTA CAT TTG GAA AAA AAT GTG AAT CAG
FTHL1  --- -TG --- --- --- --- --- --- --- --- --- --- --- --- ---
PCR5   --- -TG --- --- --- --- --- --- --- --- --- --- --- --- ---
FTHP2  --- -AG --- --- G-- --- -G- --- --- --G --- --C A-- --- T--
     555
FTH1    TCA CTA CTG GAA CTG CAC AAA CTG GCC ACT GAC AAA AAT GAC CCC
FTHL1   --- --- --- --- --- --- --- T-- --- --- --- --- --- --- ---
PCR5    --- --- --- --- --- --- --- T-- --- --- --- --- --- --- -T-
FTHP2   --- --- --- --- --- T-- --- --- --- --- --- --- --- --- -T-
      600
FTH1    CAT TTG TGT GAC TTC ATT GAG ACA CAT TAC CTG AAT GAG CAG GTG
FTHL1   --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
PCR5    --C --- --- --- --A --G --- --- A--IC-- --- --G --- --- ---
FTHP2   --C --- --- --- --A --G --- --- A--IC-- --- --G --- --- ---
                                    Insert GTG TCC TGG ATG GAT TTT
      645
FTH1    AAA GCC ATC AAA GAA TTG GGT GAC CAC GTG ACC AAC TTG CGC AAG
FTHL1   --- --- --- --- --- --- -A- --- --T --- --- --- --- -A- ---
PCR5    --- T-- --- --- T-- --- -A- --- ---A--- --- --- --- -A- ---
FTHP2   --- T-- --- --- T-- --- -A- --- ---A--- --- --- --- -A- ---
      690
FTH1    ATG GGA GCG CCC GAA TCT GGC TTG GCG GAA TAT CTC TTT GAC AAG
FTHL1   --- --- --A -T- --- --- --- --- --A --- --- --- -A- --- ---
PCR5    -C- --- --A --- A-- --- --- --- --A --- --- G-A --- --- --A
FTHP2   -C- --- --A --- A-- --- --- --- --A --- --- G-A --- --- --A
      735
FTH1    CAC ACC CTG GGA GAC AGT GAT AAT GAA AGC TAA CCTCGGGCTAATT
FTHL1   --- --- --- --- --- --- --- --- --- --- --- ------X------
PCR5    --- --- -CC C-- --- --- --C --- --G --- --- G--TA--------
FTHP2   --- --- -CC C-- --- --- --C --- --G --- --- G--TA---------
     780
FTH1  TTCCCCATAGCCGTGGGGTGACTTCCCTGGTCACCAAGGCAGTGCATGCATTG
FTHL1 ----G--------------------------------------------
PCR5  -----------A------A--C----
FTHP2 -----------A------A--C-------------------------G-
```

Figure 1. Comparison of sequences from FTH1, FTHL1, FTHP2 and PCR5. Primer sequences and common differences with the H cDNA are highlighted. The shaded areas correspond to regions of homology between PCR5 and the genomic clones.

sequences. One was unusual in that it appeared to be a hybrid of two quite different H-like sequences.

As shown in Figure 1, the sequence in PCR5 differs from the H cDNA (FTH1) in 53 of the 480 bp amplified between the primers. These differences are not randomly distributed. There are only 4 differences in the first 230 bp, but 49 in the remaining 250 bp. In the first 230 bp, the PCR5 sequence is identical except for one position to an intronless sequence, FTHL1, in a genomic clone from chromosome 11 (4). The two sequences then diverge markedly. From this point, the PCR5 sequence is now identical to a highly degenerate sequence in FTHP2 cloned from chromosome 4 (5). These similarities are seen in 28 point mutations, a 1 nt insertion and a distinctive 18 nt insertion not found so far in any other H-like sequence.

Discussion

Cloning PCR products from amplification of a multigene family with the same primer set is a convenient method for obtaining sequence of different members without conventional cloning. The ferritin H gene family is amenable to this approach. Many members are highly homologous, though some seem to represent distinct subgroups which evolved at different periods (5).

In general, the PCR approach seemed valid since several of the PCR products from genomic DNA corresponded to conventionally cloned sequences. However, our experiments indicates that some may not be true genomic sequences. The sequence in PCR5 is such an example. Its 5' end is almost identical to a highly conserved sequence from chromosome 11 while the 3' end is identical to a highly divergent sequence from chromosome 4. It seems unlikely that PCR5 represents a true genomic sequence. A more likely explanation is that it is a chimera produced during the PCR reaction from two different ferritin H sequences.

We suggest that in the mixture of amplified genomic sequences, some were not completely extended between the primer sequences (Figure 2A). In subsequent rounds, the 3' end of an incomplete sequence from one gene could anneal to the complementary region of another related sequence (Figure 2B). The incomplete sequence(s) would then be extended by copying this second sequence. In later rounds, the region between the primers would be amplified to give a chimeric sequence (Figure 2C).

We therefore suggest that the sequence in PCR5 is a hybrid formed from FTHL1 and FTHP2 or similar sequences. In support of this, we note that the ferritin sequences in FTHL1 and FTHP2 are both amplified with the primer set used here and that they share a 21 nt sequence identity in the putative fusion region. These observations suggest that it would be prudent to confirm sequences derived from PCR amplification of a mixture of related sequences by analyses of individual members obtained by other methods.

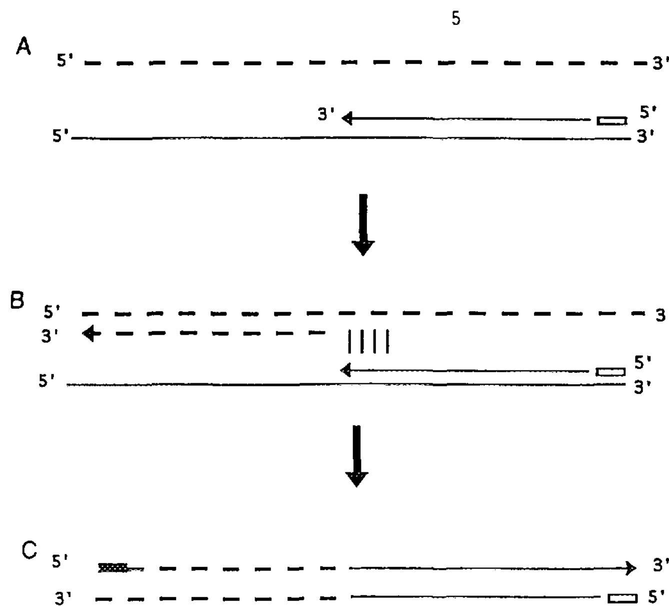

Figure 2. Formation of chimeric molecules during PCR amplification of a mixture of related sequences —— and - - -.

Acknowledgment. This work was supported by AM 17775.

References

1. Boyd, D., Vecoli, C., Belcher, D.M., Jain, S.K., Drysdale, J.W. (1985) J. Biol. Chem. 260:11755-11766.
2. Jain, S.K., Barrett, K.J., Boyd, D., Favreau, M.F., Crampton, J. and Drysdale, J.W. (1985) J. Biol. Chem. 2160:11762-11768.
3. Costanzo, F., Colombo, M., Staempfli, S., Santoro, C., Marone, M., Frank, R., Delius, H. and Cortese, R. (1986) Nucl. Acids Res. 14:721-736.
4. Papadopoulos, P., Bhavsar, D., Zappone, E., David, V., Jones, C., Worwood, M. and Drysdale, J. (1992) Cytogenet. Cell Genet.61:107-108.
5. Zheng, Hd., Bhavsar, D., and Drysdale, J. (1994) DNA Seq. (in press).