Analytical Methods

RSCPublishing

PAPER

View Article Online
View Journal | View Issue

Published on 10 July 2013. Downloaded by Chengdu Library of Chinese Academy of Science on 29/10/2013 01:05:28.

Gold nanoparticle-assisted primer walking for closing the human chromosomal gap

Cite this: *Anal. Methods*, 2013, **5**, 4746

Haikuo Li,†[a] Baochen Shi,†[c] Xueling Li,†[d] Steffen Nyegaard,[e] Ming LÜ,[b] Junhong LÜ,*[a] Xueyan Cao,[b] Qing Huang,[a] Chunhai Fan,[a] Runsheng Chen*[c] and Jun Hu*[ab]

Received 8th May 2013
Accepted 10th July 2013

DOI: 10.1039/c3ay40770e

www.rsc.org/methods

The finished sequence of the human genome still contains 260 euchromatic gaps. All the PCR-based genome walking techniques used to close gaps have common limitations, such as low efficiency and low specificity. We herein describe a strategy to solve this problem by employing gold nanoparticles (AuNPs) to improve the efficiency in primer walking amplification. We used this strategy to close a gap in human chromosome 5 containing a DNA stretch composed of the 12SAT repeat. The obtained gap sequence is highly conserved among several mammalian genomes. The demonstrated AuNP-assisted primer walking strategy is capable of effectively improving the specificity of PCR amplification and enriching the yield of target DNA fragments; it offers a new avenue for closing gaps left by current sequencing methods.

Introduction

Although the human genome sequence was generally declared "finished" in 2004,[1] the latest assembly (NCBI Build 37.2 Assemble) still contains 260 euchromatic gaps.[2,3] For example, Chromosome 5, one of the most well assembled chromosomes, has seven gaps remain.[4] To close these gaps is still a worthy goal as each missed sequence can potentially contain genes or regulatory elements.[5,6] Various polymerase chain reaction (PCR)-based primer-walking methods have been developed for acquiring the unknown genomic sequences,[7,8] including inverse PCR,[9] restriction-site dependent PCR,[10] and ligation-mediated PCR.[11] However, due to the inherent characteristics of the PCR and the comprehensive complexity of the genome, all the PCR-based genome-walking techniques have common limitations, such as low efficiency and low specificity.[12] Herein, we report a novel strategy in which gold nanoparticles (AuNPs) have been employed in PCR-mediated primer walking to improve the specific amplification for closing gaps in the human genome.

AuNPs have attracted wide attention and been proven of great utility in biological systems due to their unique physical and chemical properties.[13,14] Others and we previously found that AuNPs could remarkably improve PCR amplification in sensitivity, specificity and extension rate.[15,16] Although the definite mechanism remains unclear,[17,18] the enhancement effect was successfully used in different PCR systems for detection of viral infection[19] and identification of DNA variations.[20]

In this study, we describe and demonstrate an AuNP-assisted primer walking strategy capable of effectively improving the specificity of PCR amplification and enriching the yield of target DNA fragments. Combined with DNA sequencing, we were able to close a gap in human chromosome 5 and then analyze its sequence. We believe that AuNP-enhanced genomic walking will become a valuable approach to allow completion of the DNA sequence with the gaps left by current sequencing methods.

Materials and methods

Primer design

The initial amplification primers were designed according to the gap's flanking sequences. Both sequences were extracted from the two flanks of the gap in the current finished human genome, and known repeats were masked by RepeatMasker.[21] Primers were then constructed from the unmasked sequence by basic primer designing principles. The chosen primers were tested against the human genome by e-PCR[22] to prevent unwanted amplification products. Additional primers were then designed according to the new sequencing results for reamplification. The sequences of all the primers used below are listed in Table 1.

PCR amplification

The human genomic DNAs were obtained from BioDevTech, Beijing, China, and were purified to a concentration of

[a]*Laboratory of Physical Biology, Shanghai Institute of Applied Physics, Chinese Academy of Sciences, Shanghai 201800, China. E-mail: lujunhong@sinap.ac.cn; hujun@sinap.ac.cn; Fax: +8621 59552394; Tel: +8621 39194259*

[b]*College of Life Science and Biotechnology, Shanghai JiaoTong University, Shanghai 200240, China*

[c]*Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China. E-mail: chenrs@sun5.ibp.ac.cn*

[d]*Shanghai Center for Bioinformation Technology, Shanghai 201203, China*

[e]*Department of Molecular Biology and Genetics, Aarhus University, Aarhus, Denmark*

† These authors contributed equally to this work.

This journal is © The Royal Society of Chemistry 2013

Published on 10 July 2013. Downloaded by Chengdu Library of Chinese Academy of Science on 29/10/2013 01:05:28.

Table 1 Oligo primers used in PCR experiments and in sequencing reactions

Primer name	Primer sequence
In1	5′-ctcccaaagcactaggattacagg-3′
In2	5′-ccatcccatcagacattctcaaac-3′
W1	5′-tctctctctcaccaccttttttctc-3′
W2	5′-agcatattgtcttttcagtagcagt-3′
W3	5′-tgttcaccgtcaggaagtaaatg-3′
W4	5′-aggagcattatttggaccacttac-3′
W5	5′-ctgcatcatattcaccaacacc-3′
W6	5′-ccatgaagattaaatgagtgcc-3′
W7	5′-tcatcttcaccattgccctttc-3′
W8	5′-gttttatccctcctcaacaagc-3′

100 ng μL^{-1}. The genomic DNAs were directly used as PCR templates. The PCR primers were synthesized by Invitrogen, Shanghai, China, purified by the PAGE method, and were diluted to 10 μM with TE buffer (10 mM Tris–HCl, 1 mM EDTA, pH8.0) before usage. DNA polymerase was ExTaq™ HotStart Version from TaKaRa, Dalian, China. A DNA marker was obtained from Sangon Biotech, Shanghai, China. Colloidal gold (Φ10 nm; A_{520nm} 0.75) was bought from Sigma-Aldrich, Inc.

This gold colloid product was produced by a modified tannic acid/citrate method, containing approximately 0.01% $HAuCl_4$ suspended in 0.01% tannic acid with 0.04% trisodium citrate, 0.26 mM potassium carbonate, and 0.02% sodium azide as a preservative.

A general PCR system is as follows: 1–2 μl DNA template, 0.5 μl 10 μM primer each, 1U ExTaq™ HS (TaKaRa), 2.5 μl dNTP (2.5 mM each, TaKaRa), 2.5 μl 10× Ex PCR buffer (Mg^{2+} plus, TaKaRa), and double distilled water (supplemented into a total of 25 μl). In nanoparticle assisted PCR, 1–4 μl colloidal gold (A_{520nm} 0.75) was added to the PCR mixture with a final volume of 25 μl. A general PCR temperature cycle is as follows: 2 min pre-denaturation at 94 °C, 30 s denaturation at 94 °C, 30 s annealing (annealing temperature was set according to the primer, T_m, 2 °C lowered), elongation at 72 °C (elongation time was set according to the product length, 1 min kbp^{-1}, no less than 30 s), the total cycle number was 25–30, and finally an additional supplementary elongation at 72 °C. The PCR was done on a PTC-200 thermal cycler (MJ Research). Some trivial modifications would be made based on the general protocol in variant cases.

Sanger sequencing

PCR products were used as sequencing reaction substrates. Before directly sequencing the PCR product, a single band of the PCR product was isolated and purified using an Agarose Gel DNA Purification Kit (TaKaRa, Dalian, China). Both forward and reverse amplification primers were used as sequencing primers to ensure complete coverage of the primer overlapped regions. The Sanger sequencing reactions were performed on an ABI 3730 DNA analyzer by Invitrogen, Shanghai, China.

Results and discussion

In the present paper, we have developed an AuNP-based genomic walking strategy for closing a human chromosome 5 gap at 5q33.3 (Fig. 1), which is located in the first intron of the mRNA BC008431.1 (a transcript of the gene SGCD) that codes the delta-sarcoglycan (35 kD dystrophin-associated glycoprotein). As for other recalcitrant gaps in chromosome 5, this gap appears to be unclonable in current vector systems, presumed as a clone gap.[4] To fill these clone gaps, one alternative strategy is to directly sequence the gap-spanning PCR products. However, amplifying these gaps with conventional PCR methods is a particularly difficult problem as the sequences concerned commonly have a paucity of good oligonucleotide prime sites, which result in heavily nonspecific products. To overcome this problem, gold nanoparticles were applied to improve the specificity of PCR amplification. We found that when an optimized concentration of AuNPs was added (3.5 μl in 50 μl of reaction volume), a fairly specific target band (3.7 kb) was obtained (Fig. 2A). In contrast,

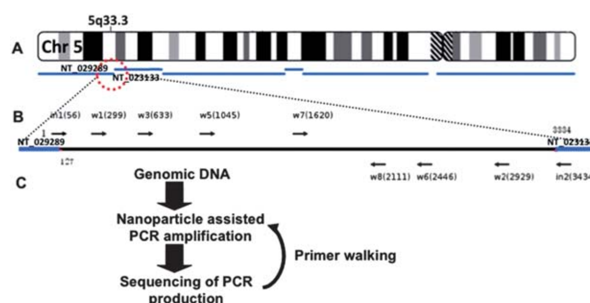

Fig. 1 (A) Schematic diagram depicting the location of gaps on chromosome 5. (B) Schematic representations of the primers used and their locations on chromosome 5. (C) AuNP-assisted primer walking strategy.

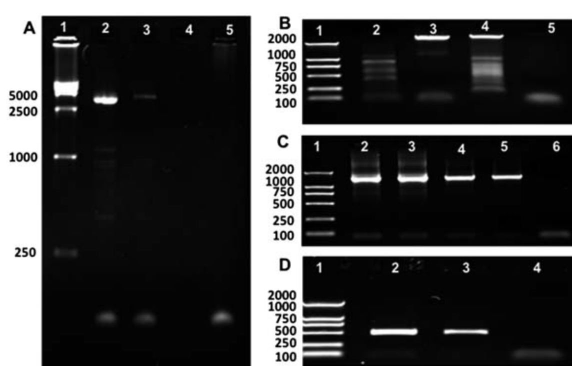

Fig. 2 Primer walking experiments for amplifying and sequencing the gap region. (A) PCR with primers In1/In2. Lane 1, 15K DNA marker; Lanes 2–4, PCR amplification with the addition of 0 μl, 3.5 μl and 4 μl of AuNPs, respectively; Lane 5, negative control. (B) The recovered PCR products (IN1IN2) were further amplified with 4 different primer pairs In1/In2 (Lane 2), W1/In2 (Lane 3), W1/W2 (Lane4), and In1/W2 (Lane 5), respectively; Lane 1, 2K DNA marker. (C) PCR with primers W5/W6. Lane 1, 2K DNA marker; Lanes 2–5, PCR amplification with the addition of 0 μl, 0 μl, 3.5 μl and 4 μl of AuNPs, respectively; Lane 6, negative control. (D) PCR with primers W7/W8. Lane 1, 2K DNA marker; Lanes 2 and 3, PCR amplification with the addition of 0 μl and 3.5 μl of AuNPs, respectively; Lane 4, negative control. Primers are reported in Table 1.

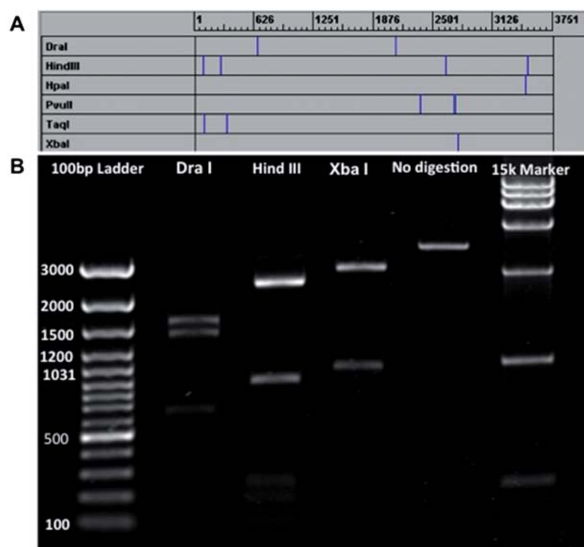

Fig. 3 Gap sequence assembly validation by multiple-complete-digest restriction fragment analysis. (A) Predicted fragment sizes and the location of six single-restriction enzyme recognition sites within the gap sequence. (B) The experimental restriction digests and agarose gel electrophoresis results. Lanes labeled with the 100 bp ladder and 15K marker identify size markers.

without AuNP addition, there were several visible nonspecific bands. It should be noted that, consistent with previous reports,[15,18] high concentration of AuNPs could inhibit the amplification reaction.

The specific PCR products (IN1IN2) were recovered and sequenced with a direct sequencing approach using the same primers in the PCR reaction. As expected, the sequencing quality of the PCR products with AuNPs was greatly improved, which made possible for subsequent PCR walking. However, we failed to get the target segment when the PCR production was cloned (pGEM-T vector system, Promega) and sequenced, which is confirmed to be a clone gap.

It should be mentioned that, although the total PCR products with AuNPs were significantly decreased (the weak band in the gel, as shown in Lane 3 in Fig. 2A), the sequencing quality (base quality values scored by phrap software) was improved around 20% than that of PCR products without AuNPs, indicating that the AuNP-enhanced specificity produced highly pure PCR products. After obtaining partial sequences of the gap region using the IN1 and IN2 primer pair, we further performed primer walking using the IN1IN2 product as the template and specific primers close to the end of the known sequence obtained (Table 1, Fig. 1 and Fig. 2B–D). As shown in Fig. 2B–D, the AuNP-assisted primer walking worked well in all the PCR reactions and could make it possible to obtain PCR products with high purity for sequencing. Finally, 37 sequence reads were obtained and assembled to a consensus sequence with 3207nt (GenBank accession DQ841277) using the phrap program (http://www.phrap.org/).

To test the fidelity of the gap sequence assembly, we used a restriction digest fingerprint matching technique,[23] which has been extensively used as the basis for physical map assembly. The pattern of fragment masses resulting from a restriction digest of the source DNA can be readily determined by gel electrophoresis with a precision of ±1%. Because the cleavage sites for restriction enzymes are well established, using the previously described dynamic programming algorithm, it is

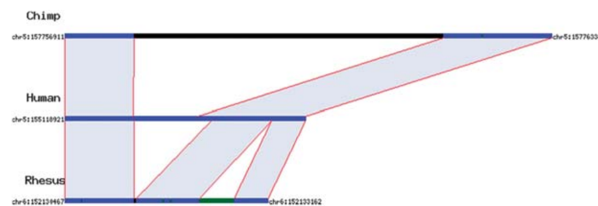

Fig. 5 Sequence conservation of the gap sequence among the human, chimp and rhesus genome.

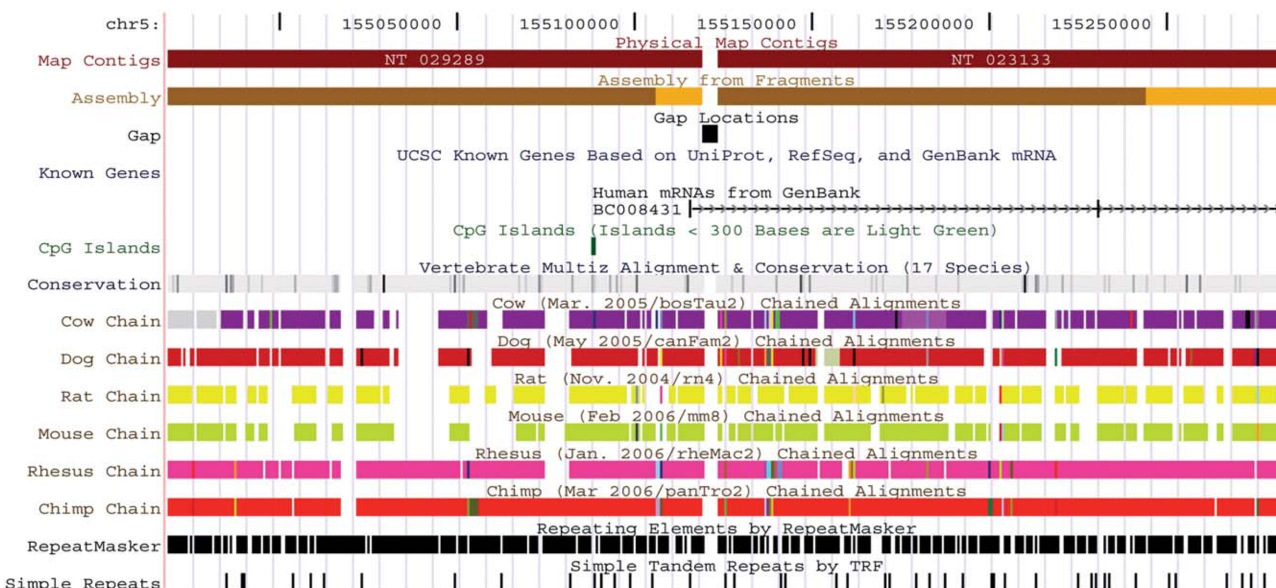

Fig. 4 Blast analysis of the gap sequence in human chromosome 5 with the corresponding regions in other six mammalian genomes.

Published on 10 July 2013. Downloaded by Chengdu Library of Chinese Academy of Science on 29/10/2013 01:05:28.

 This journal is © The Royal Society of Chemistry 2013

Published on 10 July 2013. Downloaded by Chengdu Library of Chinese Academy of Science on 29/10/2013 01:05:28.

trivial to electronically generate a set of predicted fragment masses from the finished sequence. In our experiment, three restriction enzymes with 6 bp recognition sites (*Dra* I, *Hind*III and *Xba* I) were used to completely and independently digest the PCR products of the gap (Fig. 3B). The fragments digested by each enzyme were compared between results obtained from agarose gel electrophoresis and the predicted pattern of restriction fragments (Fig. 3A). The experimental data agree well with the predicted fragment mobilities, indicating an accurate assembly of the target sequence.

The filled gap sequence has low GC content (39.13%), and seems to extend the first intron of the mRNA BC008431 to 114 621 bp. Similar to its flanking sequences, the closed gap is rich in repeat sequences (tandem repeated 12 bp unit, 12SAT), and is conserved across seven mammalian genomes (cow, dog, rat, mouse, chimp, rhesus and human) (Fig. 4), especially among the three primates (human, chimp and rhesus). Since the corresponding region of the chimp genome also contains an estimated 4.1k gap, the sequence we closed in human chromosomes is also helpful to fill a chimp gap. However, the corresponding region in the Rhesus genome contains 465 bp insert sequence (Fig. 5).

Conclusions

This paper describes an alternative strategy for amplification and sequencing of a gap on the human chromosome 5 by primer walking combined with AuNP-assisted PCR. Sequence analysis of the gap region reveals a DNA stretch composed of the 12SAT repeat. This repeat could account for the inability to clone this region. Although several mechanisms have been proposed to explain the enhancement of gold nanoparticles in PCR amplification, such as condensation of PCR reactions on the gold surfaces,[24] improving heat transfer,[16] modulation of DNA polymerases,[17,18] and SSB-like mechanisms,[15] none of them can account for all the observed phenomena. Moreover, due to the complexity of human genome gaps and PCR systems, the efficiency enhancement of gold nanoparticles in primer walking amplification should be evaluated case-by-case. Given its fast, accurate and cost-effective merit, the potential and high applicability in other gaps make the demonstrated method an interesting alternative to the well-established PCR- and cloning based genome-walking technique.

Notes and references

1 International Human Genome Sequencing Consortium, *Nature*, 2004, **431**, 931–945.
2 E. E. Eichler, R. A. Clark and X. She, *Nat. Rev. Genet.*, 2004, **5**, 345–354.
3 M. Garbre, M. C. Zody, H. M. Arachchi, A. Berlin, S. Gnerre, M. Lisa, L. M. Green, N. Lennon and C. Nusbaum, *Genome Biol.*, 2009, **10**, R60.
4 J. Schmutz, J. Martin, A. Terry, O. Couronne, J. Grimwood, S. Lowry, L. A. Gordon, D. Scott, G. Xie, W. Huang, U. Hellsten, M. Tran-Gyamfi, X. She, S. Prabhakar, A. Aerts, M. Altherr, E. Bajorek, S. Black, E. Branscomb, C. Caoile, J. F. Challacombe, Y. M. Chan, M. Denys, J. C. Detter, J. Escobar, D. Flowers, D. Fotopulos, T. Glavina, M. Gomez, E. Gonzales, D. Goodstein, I. Grigoriev, M. Groza, N. Hammon, T. Hawkins, L. Haydu, S. Israni, J. Jett, K. Kadner, H. Kimball, A. Kobayashi, F. Lopez, Y. Lou, D. Martinez, C. Medina, J. Morgan, R. Nandkeshwar, J. P. Noonan, S. Pitluck, M. Pollard, P. Predki, J. Priest, L. Ramirez, J. Retterer, A. Rodriguez, S. Rogers, A. Salamov, A. Salazar, N. Thayer, H. Tice, M. Tsai, A. Ustaszewska, N. Vo, J. Wheeler, K. Wu, J. Yang, M. Dickson, J. F. Cheng, E. E. Eichler, A. Olsen, L. A. Pennacchio, D. S. Rokhsar, P. Richardson, S. M. Lucas, R. M. Myers and E. M. Rubin, *Nature*, 2004, **431**, 268–274.
5 N. Kouprina, S. H. Leem, G. Solomon, A. Ly, M. Koriabine, J. Otstot, E. Pak, A. Dutra, S. Zhao, J. C. Barrett and V. Larionov, *EMBO Rep.*, 2003, **4**, 257–262.
6 S. H. Leem, N. Kouprina, H. Grinwood, J. H. Kim, M. Mullokandov, Y. H. Yoon, J. Y. Chae, J. Morgan, S. Lucas, P. Richardson, C. Detter, T. Glavina, E. Rubin, J. C. Barrett and V. Larionov, *Genome Res.*, 2004, **14**, 239–246.
7 J. Kieleczawa, J. J. Dunn and F. W. Studier, *Science*, 1992, **258**, 1787–1791.
8 Y. Tonooka and M. Fujishima, *Appl. Microbiol. Biotechnol.*, 2009, **85**, 37–42.
9 T. Uchiyama and K. Watanabe, *Biotechniques*, 2006, **41**, 183–185.
10 Y. Jiang, J. Pei, X. Song and W. Shao, *DNA Res.*, 2007, **14**, 283–290.
11 B. B. Orcheski and T. M. Davis, *BMC Res. Notes*, 2010, **3**, 61.
12 J. H. Bae and J. H. Sohn, *Anal. Biochem.*, 2010, **398**, 112–116.
13 C. J. Murphy, A. M. Gole, J. W. Stone, P. N. Sisco, A. M. Alkilany, E. C. Goldsmith and S. C. Baxter, *Acc. Chem. Res.*, 2008, **41**, 1721–1730.
14 D. A. Giljohann, D. S. Seferos, W. L. Daniel, M. D. Massich, P. C. Patel and C. A. Mirkin, *Angew. Chem., Int. Ed.*, 2010, **49**, 3280–3294.
15 H. Li, J. Huang, J. Lv, X. Zhang, Z. Zhang, C. Fan and J. Hu, *Angew. Chem., Int. Ed.*, 2005, **44**, 5100–5103.
16 M. Li, Y. C. Lin, C. C. Wu and H. S. Liu, *Nucleic Acids Res.*, 2005, **33**, e184.
17 L. J. Mi, H. P. Zhu, X. D. Zhang, J. Hu and C. Fan, *Chin. Sci. Bull.*, 2007, **52**, 2345.
18 L. J. Mi, Y. Wen, D. Pan, Y. Wang, C. Fan and J. Hu, *Small*, 2009, **5**, 2597–2600.
19 S. H. Huang, T. C. Yang, M. H. Tsai, I. S. Tsai, H. C. Lu, P. H. Chuang, L. Wan, Y. J. Lin, C. H. Lai and C. W. Lin, *Nanotechnology*, 2008, **19**, 405101.
20 P. Chen, D. Pan, C. Fan, J. Chen, K. Huang, D. Wang, H. Zhang, Y. Li, G. Feng, P. Liang, L. He and Y. Shi, *Nat. Nanotechnol.*, 2011, **6**, 639–644.
21 http://www.repeatmasker.org/faq.html.
22 K. Rotmistrovsky, W. Jang and G. D. Schuler, *Nucleic Acids Res.*, 2004, **32**, W108–W112.
23 E. C. Rouchka and D. J. States, *Proc Int Conf Intell Syst Mol Biol*, 1998, **6**, 140–214.
24 X. Cao, X. Shi, W. Yang, X. Zhang, C. Fan and J. Hu, *Analyst*, 2009, **134**, 87–92.