



A new approach using multiplex long accurate PCR and yeast artificial chromosomes for bacterial chromosome mapping and sequencing.

A Sorokin, A Lapidus, V Capuano, et al.

Genome Res. 1996 6: 448-453

Access the most recent version at doi:[10.1101/gr.6.5.448](https://doi.org/10.1101/gr.6.5.448)

References

This article cites 14 articles, 5 of which can be accessed free at:
<http://genome.cshlp.org/content/6/5/448.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](#).

A horizontal banner advertisement with a teal background. On the left, the text reads 'CRISPR and RNAi Genetic Screening. Your new superpower.' in white. In the center, there is a white-bordered box containing the words 'LEARN MORE' in black. On the right, there is a photograph of a woman wearing a red mask and a red cape, and the Cellecta logo, which consists of a green molecular structure and the word 'CELLECTA' in white.

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

Copyright © Cold Spring Harbor Laboratory Press

GENOME METHODS

A New Approach Using Multiplex Long Accurate PCR and Yeast Artificial Chromosomes for Bacterial Chromosome Mapping and Sequencing

Alexei Sorokin,¹ Alla Lapidus, Veronique Capuano, Nathalie Galleron, Petar Pujic, and S. Dusko Ehrlich

Laboratoire de Génétique Microbienne, Institut National de la Recherche Agronomique, Domaine de Vilvert, 78352 Jouy en Josas CEDEX, France

An efficient approach for structural studies on bacterial chromosomes is presented. It is based on high-resolution PCR map construction by using a multiplex long accurate PCR (MLA PCR) protocol and a YAC clone carrying the region to be mapped as indicator. The high-resolution PCR map of the *Bacillus subtilis* *rrnB*–*dnaB* region is presented as an example. Data are also presented on the use of DNA generated by LA PCR for sequencing; they are relevant to LA PCR induced mutations and justify the application of such mapping for sequencing long stretches of bacterial chromosomes.

DNA cloning in yeast artificial chromosomes (YACs, Burke et al. 1987) has allowed the establishment of ordered collections of genome segments of organisms ranging from bacteria to man (Azevedo et al. 1993; Cohen et al. 1993). An attractive feature of such collections is that they are composed of a relatively small number of clones, because of the capacity of YACs to carry inserts of large size. However, large size may be a drawback for direct sequence analysis, and the use of YACs in sequencing projects has been limited to the construction of ordered λ or cosmid-based collections, by subcloning (Marshall 1995). We decided to follow, within the *Bacillus subtilis* Genome Sequencing Project (Kunst et al. 1995; Ogasawara et al. 1995), a different strategy for YAC sequencing. The strategy is based on a two-stage procedure: (1) shotgun sequencing of short YAC fragments cloned in phage M13, until a convenient number of contigs is obtained; and (2) ordering the contigs by combinatorial PCR, using as primers oligonucleotides complementary to the ends of this contig, and closing the gaps between contigs by sequencing the PCR products. A 104-kb insert of YAC 15-6B, from the ordered collection of the *B. subtilis* chromosome segments (Azevedo et al.

1993), was sequenced in this way (Sorokin et al. 1996). The advent of long accurate PCR [(LA PCR) Barnes 1994; Cheng et al. 1994] greatly improved this approach. The ordering of the contigs generated a map of the region, which we call the PCR map. This mapping is the inherent feature of the strategy. We describe here an efficient LA PCR-based strategy to order contigs and generate material for further sequencing, and discuss the quality of the sequence obtained with LA PCR products.

RESULTS AND DISCUSSION

Constructing of a Minimal PCR Map of the *rrnB*–*dnaB* Region of the *B. subtilis* Chromosome

We used YAC 15-132, which covers most of the region between the *dnaB* and *rrnB* loci of the *B. subtilis* chromosome (Azevedo et al. 1993) and carries an insert of ~170 kb, as deduced from pulse field gel electrophoresis (PFGE). PFGE-purified DNA of this YAC was cloned in phage M13, and 500 clones were sequenced randomly, yielding 42 representative contigs, containing at least two independent gel readings.

Contig ordering is based on simultaneous use of multiple primers (Fig. 1, top). As depicted in Figure 1, concomitant use of six primers complementary to the ends of three contigs yields two

¹Corresponding author.
E-MAIL sorokine@biotec.jouy.inra.fr; FAX 33-1 34652521.

YAC SEQUENCING

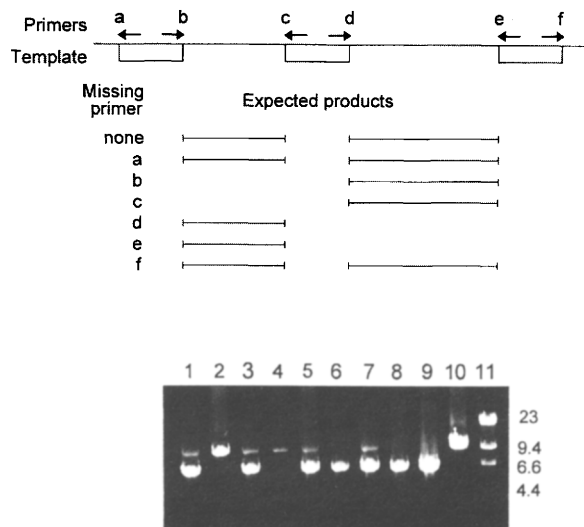


Figure 1 Ordering of contigs by MLA PCR. (Top) Schematic representation of the method. Template DNA is drawn as a line, contigs are represented by open boxes, arrows indicate the position and orientation of the primers, which are identified by letters *a–f*, horizontal bars indicate the region of the template homologous to the products expected to form with different primer mixtures. (Bottom) Electrophoretic analysis of the products obtained with seven primers complementary to the ends of contigs established during sequencing of YAC 15-132. The oligonucleotides in the mixture were designated AB6, AH1, BA1, BB1, AH2, AF8, and BA7. The LA PCR products obtained with the mixture of all primers are in lane 1. The missing primers are as follows: (Lane 2) AB6; (lane 3) AH1; (lane 4) BA1; (lane 5) BB1; (lane 6) AH2; (lane 7) AF8; (lane 8) BA7. Products obtained with the primer couples AB6/BA1 and AH2/BA7 are in lanes 9 and 10, respectively. (Lane 11) Size markers (λ DNA cleaved with *Hind*III). Sizes (in kb) are indicated at right. Five microliters of a 50- μ l reaction mixture was loaded onto the gels.

PCR products (primers *b* and *e* are supposed to be too far to give a product). The primers required for synthesis of each product can be identified unambiguously by withdrawing one primer at a time from the mixture, and the relative order and orientation of contigs can thus be established. Determination of product size gives the distance between the contigs. An experiment with seven primers illustrates this approach (Fig. 1, bottom). Primers missing in lanes 2 and 4 are required to synthesize the shorter product, whereas primers missing in lanes 8 and 6 are required for the longer product. This assignment is confirmed by the use of the relevant primer couples in indi-

vidual reactions (lanes 9 and 10). Up to 16 primers were used successfully together (not shown).

The main advantage of this method, which we designated multiplex LA PCR (MLA PCR), as compared to the use of primer couples in separate reactions, is that it requires a lower number of reactions to order contigs. The difference is considerable when contig number is high. For instance, only 17 simultaneous reactions are needed to order eight contigs by MLA PCR, as opposed to 120 reactions with primer couples. Furthermore, the DNA products (one or several) obtained in each MLA PCR reaction can be purified and used for further analysis, whereas most of the reactions with primer couples do not give products. A limitation of MLA PCR is that the synthesis of one product may interfere with that of another (Fig. 1, bottom; cf. the amount of the longer product in lanes 2 and 4 with that in lanes 3, 5, and 7). This can lead to failure to order some of the contigs and may limit the number of primers that can be used in a mixture. In our mapping experiments the 16 primers used together were expected to yield no more than two products, which were detected in all cases. However, simultaneous use of 22 primers, chosen to allow synthesis of 11 products covering the mapped region, yielded only a single product, shortest in size (not shown). Further study of MLA PCR conditions would be required to clarify this phenomenon.

The mapping is an iterative process aimed at establishing relative primer orientations and distances. Each iteration represents two MLA PCR experiments, one destined to measure the distance between two or more primers in a mixture and the second to identify the primers that gave the PCR products. After each iteration step the map is updated graphically or analytically, as described in Methods.

To apply this mapping strategy to YAC 15-132, 32 representative contigs were chosen and primers homologous to their ends were synthesized. A few data on the sequencing of this region were reported earlier (for a compilation, see Anagnostopoulos et al. 1993), which together with our sequenced tags contained 46,000 bp. The unknown sequence thus contained 129,000 bp, and the average distance between representative contigs was therefore 4 kb (129/32). Iteration of MLA PCR mapping resulted in the ordering of 23 contigs, with 9 remaining unpositioned. We call a minimal PCR map one of the minimal sets of

SOROKIN ET AL.

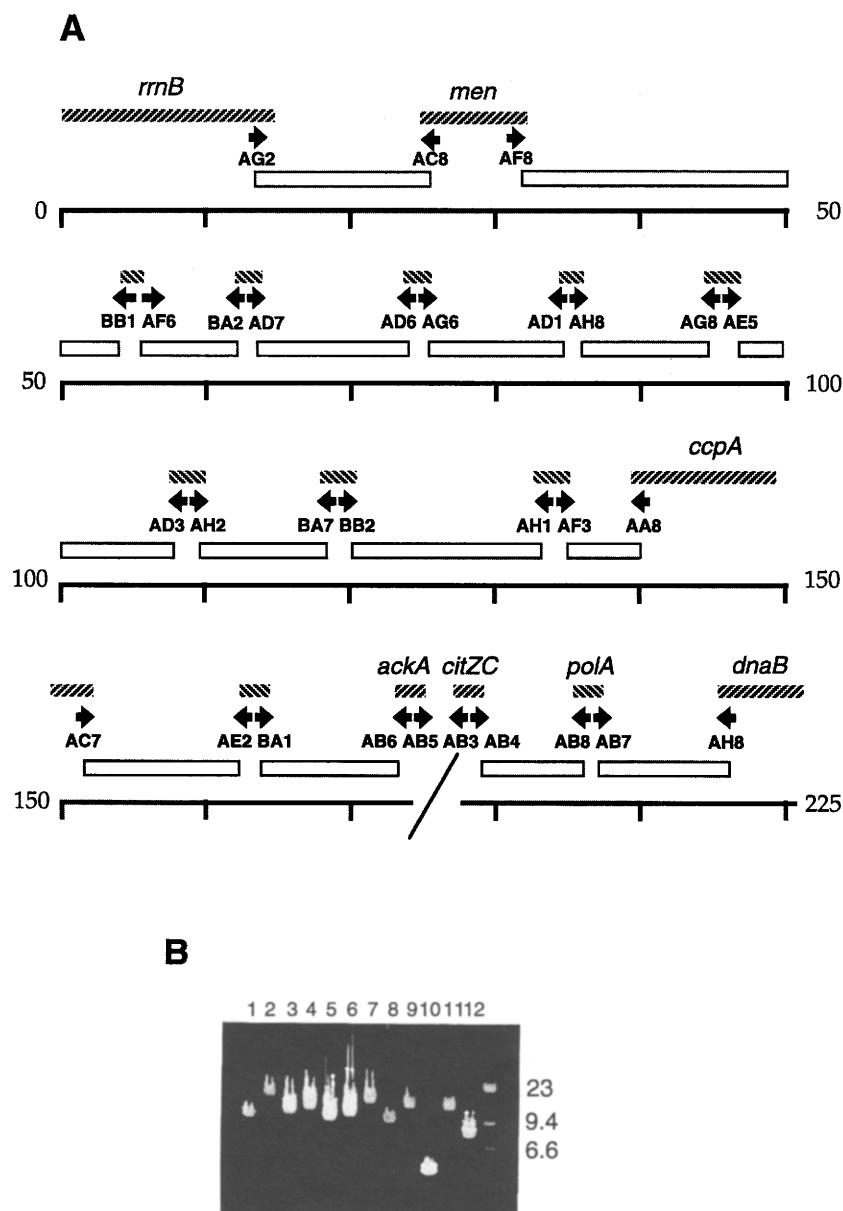


Figure 2 PCR map of the *rrnB*–*dnaB* region of the *B. subtilis* 168 chromosome. (A) Only the minimal PCR map described in the text is shown. YAC 15-132 covers the region between the *rrnB* and *ackA* genetic loci. The map of the region between *citZC* and *dnaB* was obtained by using primers corresponding to previously sequenced loci: *citZC*, *polA*, and *dnaB*. The area between *ackA* and *citZC* was mapped by using YAC 11-105 (see the text for details). Hatched rectangles represent sequenced areas. White rectangles correspond to the LA PCR fragments. Arrows with names designate primers. Sequenced loci are designated by the names of the genes according to the *B. subtilis* genetic map (Anagnostopoulos et al. 1993). Scale is in kb. (B) LA PCR amplification of the YAC 15-132 region. The primers used for the amplification are as follows: (1) AG2 + AC8; (2) AF8 + BB1; (3) AF6 + BA2; (4) AD7 + AD6; (5) AG6 + AD1; (6) AH8 + AG8; (7) AE5 + AD3; (8) AH2 + BA7; (9) BB2 + AH1; (10) AF3 + AA8; (11) AC7 + AE2; (12) BA1 + AB6; (13) λ /HindIII size markers (in kb).

contigs and corresponding primers that allows amplification of the whole mapped region. Such a map for YAC 15-132, comprising 13 contigs of known sequence ranging in size from 472 to 15,558 bp, is shown in Figure 2A. The gaps between the contigs range from 6 to 18 kb, and the corresponding LA PCR products can readily be obtained by using appropriate primer couples (Fig. 2B). Ten other contigs of known sequence were also positioned on the map (not shown), and therefore different combinations of them can be chosen to obtain variants of the minimal map. We used previously determined sequences of the genes in the neighborhood of the YAC 15-132 region to extend the PCR map to *dnaB* (Fig. 2A). The gap (~30 kb) between *ackA* (the last YAC marker) and *citZC* is bigger than the longest fragment that we were able to synthesize by LA PCR (~18 kb). We therefore used another YAC from our collection, 11-105, in the same way as 15-132, to generate contigs within the gap. This allowed completion of the PCR map of the *rrnB*–*dnaB* region (data not shown).

A parameter that can be used to characterize a map is its resolution. This can be defined in our case as the size of the mapped region (in kb), divided by the number of mapped contigs. A minimal map is then the one with the lowest resolution. The resolution of the YAC 15-132 complete map was 7.6 kb (175 kb/23) and that of the minimal map was 13.4 kb (175 kb/13). It should be noted that the highest resolution map (and eventually the nucleotide sequence) can be used for confirmation of lower resolution maps.

To optimize the speed of the MLA PCR mapping and minimize its cost, the number of con-

tigs should be close to that which is necessary and sufficient. In the case of YAC 15-132, to map 129 kb of unsequenced region, 32 contigs were used, but only 23 were required to complete the map. The average distance between the mapped contigs was thus ~5 kb, which is about four times less than the maximal length of the LA PCR product that we obtained (~18 kb). This ratio could provide guidance in future mapping experiments.

Use of LA PCR products for Sequencing

The use of LA PCR products for sequencing was tested during sequencing of the region between primers AG2-AC8 and AF8-AG8 containing 67,600 bp (Fig. 2A).

At least 10 independent LA PCR reactions were carried out for each region, using *B. subtilis* chromosomal DNA as template. All products were mixed to decrease the proportion of molecules with incorrect sequence, which could accumulate if a mistake occurred during early amplification cycles in one of the reactions. The products were used either for construction of random banks in plasmid pSGMU2 (Errington 1986) and the M13mp18 phage or for direct sequencing by synthetic primers. A total of 695 sequencing reactions were carried out with M13 or plasmid clones giving a total of ~288,000 bases and establishing a consensus sequence of 67.6 kb. We detected 42 point mutations in the cloned LA PCR products. The error rate generated by LA PCR is thus ~1 in 7000 bp (42/288,000), which is acceptable for genome sequencing projects. A similar result was obtained in the sequencing of a 35-kb region of another YAC 10-9 (V. Capuano, N. Galleron, P. Pujic, A. Sorokin, and S.D. Erlich, in prep.), where ~200,000 bases were read in 500 sequencing reactions and 21 mutations were detected. The correct sequence was determined in these cases by multiple sequencing of independent LA PCR-generated clones and/or direct sequencing of PCR products. We detected that in 30 cases the mutation change was A → G or T → C. This bias suggests that the mutations are attributable to the *rTth* polymerase mistakes rather than chemical modifications such as deamination of C residues during LA PCR, which would result in a bias in the opposite direction (C → T). The shift to higher GC content of product compared to template might be explained by preference of *rTth* or *Vent* polymerase for G or C bases.

CONCLUSIONS

The results presented here indicate that a high-resolution map and the sequence of DNA segments >100 kb from bacterial chromosomes can be established efficiently by a strategy that combines shotgun sequencing and MLA PCR. The strategy is particularly useful for sequencing genomes that are difficult to subclone in *Escherichia coli* as large segments, such as that of *B. subtilis* (Azevedo et al. 1993) and at least certain other AT-rich Gram-positive bacteria (M.-C. Chopin and S.D. Ehrlich, unpubl.). The method could easily be applied to segments of small genomes generated by rare cutting enzymes and purified by PFGE. Several further improvements of the strategy can be predicted: (1) The length of routine LA PCR products, which was ~18 kb in our hands, could be increased; (2) the multiplexing algorithm could be further optimized and the analysis of the results automatized; and (3) the MLA PCR conditions could be adjusted, to allow use of a higher number of primers in the mixtures. We believe that such improvements would allow extension of the mapping and subsequent sequencing to regions of megabase size, and even entire small genomes.

METHODS

Strains and Growth Conditions

Yeast cells containing artificial chromosomes were grown as described (Azevedo et al. 1993) *E. coli* JJC 128F', *araD139* Δ (*ara-leu*)7696 *galE15 galK16* Δ (*lac*)X74 *hsdr*⁻ *hsdm*⁺ *Str*^R F'[*lac*^R Δ (*lacZ*)M15 *traD36*], which is reproducibly transformed by electroporation of M13 DNA with an efficiency of 5×10^8 PFU/ μ g, was used in cloning experiments. Electrocompetent cells of JJ128F' were prepared according to the protocol described by Dower et al. (1988) and stored at -80°C. They can be used for 6 months without detectable loss of the transforming activity. *B. subtilis* 168 strain used for the chromosomal DNA preparation was supplied by C. Anagnostopoulos (this laboratory).

Isolation of Bacterial Chromosomal DNA

Overnight cultures of 25 ml in LB (Sambrook et al. 1989) were harvested and treated with lysozyme (10 mg/ml in 50 mM Tris-HCl at pH 8.0, 50 mM EDTA, 25% sucrose). SDS was added to a final concentration of 0.5% and proteinase K to 100 μ g/ml. Incubation was carried out at 50°C for 4 hr, and extraction was done twice with water-saturated phenol-chloroform (1:1 volume) at pH 8.0. After precipitation with two volumes of ethanol in 0.3 M NaOAc (pH 4.8), the DNA was removed with a glass rod and washed in 70% ethanol. It was stored in water at 100 μ g/ml.

SOROKIN ET AL.

Isolation of YAC DNA

Plugs containing yeast chromosomes were made as described by Anand et al. (1989). Electrophoresis was performed in 1.2% agarose gels in $0.3 \times$ TBE at 10 V/cm at 18°C ($1 \times$ TBE = 90 mM Tris borate/2 mM EDTA). The switching interval was from 0.3 to 6 sec in forward migration and from 0.1 to 2 sec in reverse migration for 20 hr. After separation, YAC DNA was electroeluted in a dialysis membrane (Amicon) for 4 hr at 10 V/cm. The solution was concentrated five times by butanol treatment, and the DNA was precipitated with ethanol. Phenol–chloroform treatment was usually necessary to make DNA digestible by restriction enzymes. This protocol allowed us to obtain ~1–10 μ g of YAC DNA from 500 ml of yeast cell culture. The contamination by yeast chromosomal DNA was estimated to be between 20% and 50%.

Construction of an M13 Bank

DNA (0.5–10 μ g) was partially digested by one of the following restriction enzymes: *AclI*; *TaqI*; *HpaII*; *HinPI*; or *Sau3A*. The digested DNA was separated in an agarose gel, and segments between 500 and 1500 bp were purified as described (Sorokin et al. 1993). The purified DNA was ligated into 50 ng of *SmaI*-, *AccI*-, or *BamHI*-cut and dephosphorylated M13mp19 vector in 20 μ l of ligase buffer (Boehringer), precipitated by isopropanol using tRNA or glycogene as a carrier, rinsed with 70% ethanol, dried, dissolved in deionized water, and used for electroporation. The yield of phage plaques varied from 10^5 to 10^7 per 1 μ g of M13 DNA with a ratio of white to blue clones from 1 to 10, respectively.

Selection of M13 Clones for Random Sequencing

White M13 plaques were propagated for 5–15 hr in 48-well plates in $2 \times$ YT medium on JJC128F' cells as a host. After centrifugation in a GPKR Beckman centrifuge, 150 μ l of the phage supernatants was distributed in 96-well plates by use of a BIOMEC 1000 Laboratory work station (Beckman). Fifty microliters of the supernatants was used to prepare filters corresponding to the stock plates. PFGE-purified YAC DNA labeled with 32 P by a random-prime labeling kit (Boehringer) was hybridized with the phage DNA on the filters. Only hybridizing clones were taken for sequencing.

Sequencing

Single-stranded DNA of M13 phages was prepared as described earlier (Sorokin et al. 1993). For reverse sequencing, we used double-stranded DNA of the inserts prepared by PCR with biotinylated primers and magnetic beads, according to the supplier's protocol (DYNAL). Plasmid DNA for sequencing was prepared as described (Sorokin et al. 1995). PCR products used for sequencing with dye terminators were purified by the Wizard PCR* Preps kit (Promega) or agarose gel electrophoresis. Direct and reverse PCR sequencing was performed by use of Applied Biosystems PRISM direct or reverse sequencing kit on the Perkin

Elmer 9600 thermal cyclor or the Catalyst station of Applied Biosystems.

Oligonucleotide Synthesis and Standard PCR

Oligonucleotides were synthesized in the DNA synthesizer Oligo 1000 from Beckman. Standard PCR was performed by using M13 DNA or supernatants and *B. subtilis* or yeast chromosomal DNA under conditions described previously (Sorokin et al. 1993). Primers used for LA PCR were 20–22-mers, chosen to contain 12 GC bases. The sequences of primers corresponding to the minimal PCR map shown in Figure 2A are as follows:

AA8-₅-TTGGATCACGTGAGCCAGCAGA₃; AB3-₅-AAGACCGCGTGTGCTGTCATA₃;
 AB4-₅-GGATGGAACGAAGCGGCTGAT₃; AB5-₅-GCGGCGAGGAAGCTTTCATCA₃;
 AB6-₅-TTGGCAGCTGCTGCAACAGTGCCT₃; AB7-₅-GCATCAGGCCATCTTGGTAC₃;
 AB8-₅-CAGCTTGGCAGCCATGTCAATC₃; AC7-₅-CTTCCGGACGTGGTTGCTTCA₃;
 AC8-₅-GTCCAAATGGCCTGCCTTACC₃; AD1-₅-CTTGACGCCATTGTGGCCTGT₃;
 AD3-₅-TGGATGCAGGAGCGCTATCTC₃; AD6-₅-AAGAACCAGCCGATTGGACG₃;
 AD7-₅-AGCCCAAGCTCCACCCACAAA₃; AE2-₅-CCAGCCTTCTCCAATGAGCCA₃;
 AE5-₅-GCTGGCAGCATTAGAAGCGCA₃; AF3-₅-CGTGTACTATGTCTCCGAAGCG₃;
 AF6-₅-ATTCGTTTCGCTGTGCCGTAG₃; AF8-₅-ATGCATTCTCTCGGCCCGGT₃;
 AG2-₅-GATAGCTTGCCTCTGCTGT₃; AG6-₅-GAGAGCTGGCGGTTTGGCTTT₃;
 AG8-₅-GGTGGGACTTTGAATCTCGC₃; AH1-₅-GAGCGATTCTCGTCTGAAGG₃;

MLA PCR

The conditions for the MLA PCR reaction were as follows: 20 mM tricine (pH 8.7); 85 mM KOAc; 1 mM Mg(OAc)₂; 8% glycerol; 2% dimethylsulfoxide; 0.2 mM each dNTP; 0.2 μ M each primer; 0.1 μ g of *B. subtilis* chromosomal DNA; 2 units of *rTth* (Perkin Elmer); and 0.05 units *Vent* polymerase (Biolabs). The final reaction volume was 50 μ l. Cycling conditions: 94°C for 5 min; 12 cycles of 10-sec melting at 94°C and 12 min annealing–polymerization–repair at 68°C, followed by 24 cycles with extension times increased by 15 sec for each cycle.

Mapping Formalization

Sequenced contigs were ordered graphically by redrawing a new version of the PCR map when new distances between primers were established. The procedure can be made analytically, which allows for easy automatization. Let X_{ij} be the coordinate of primer j in the system of coordinates where primer i is at point zero. We consider that these coordinates are >0 in the $5' \rightarrow 3'$ and <0 in the opposite direction from primer i . A formalization of this mapping would be $X_{ik} = X_{ij} - [(X_{ij})(X_{ji})/|X_{ij}||X_{ji}|]X_{jk}$. The mapping starts from obtaining some $X_{kj} = X_{jk} > 0$ in a successful PCR reaction and is thus an iterative process aimed at establishing all X_{ij} , $i, j = 1, \dots, N$, where N is the number of primers.

Computing

R. Staden's *xbap* programme (Dear and Staden 1991), version 9.0, was used for gel assembling and consensus sequence generating.

ACKNOWLEDGMENTS

This work was supported in part by grants from Groupe de Recherches et d'Etudes sur les Génomes (GREG), the French Genome Project, (décision 21) and the European Union (BIO2-CT93-0272). The sequence data described in this paper have been submitted to the DDBJ/EMBL/GenBank data libraries under accession nos. D83243 and D83244.

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

- Anagnostopoulos, C., P.J. Piggot, and J.A. Hoch. 1993. The genetic map of *Bacillus subtilis* 168. In *Bacillus subtilis and other gram-positive bacteria* (ed. A.L. Sonenshein, J.A. Hoch, and R. Losick), pp. 425–461. American Society for Microbiology, Washington, D.C.
- Anand, R., A. Villasante, and C. Tyler-Smith. 1989. Construction of yeast artificial chromosome libraries with large inserts using fractionation by pulsed-field electrophoresis. *Nucleic Acids Res.* **17**: 3425–3433.
- Azevedo, V., E. Alvarez, E. Zumstein, G. Damiani, V. Sgaramella, S.D. Ehrlich, and P. Serror. 1993. An ordered collection of *Bacillus subtilis* DNA segments cloned in yeast artificial chromosomes. *Proc. Natl. Acad. Sci.* **90**: 6047–6051.
- Barnes, W.M. 1994. PCR amplification of up to 35-kb DNA with high fidelity and high yield from λ bacteriophage templates. *Proc. Natl. Acad. Sci.* **91**: 2216–2220.
- Burke, D.T., G.F. Carle, and M.W. Olson. 1987. Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. *Science* **236**: 806–812.
- Cheng, S., C. Fockler, W.M. Barnes, and R. Higuchi. 1994. Effective amplification of long targets from cloned inserts and human genomic DNA. *Proc. Natl. Acad. Sci.* **91**: 5695–5699.
- Cohen, D., I. Chumakov, and J. Weissenbach. 1993. A first-generation physical map of the human genome. *Nature* **366**: 698–701.
- Dear, S. and R. Staden. 1991. A sequence assembly and editing for efficient management of large projects. *Nucleic Acids Res.* **19**: 3907–3911.
- Dower, W.J., J.F. Miller, and C.W. Ragsdale. 1988. High-efficiency transformation of *E. coli* by high-voltage electroporation. *Nucleic Acids Res.* **16**: 6127–6145.
- Errington, J. 1986. A general method for fusion of the *Escherichia coli lacZ* gene to chromosomal genes in *Bacillus subtilis*. *J. Gen. Microbiol.* **132**: 2953–2961.
- Kunst, F., A. Vassaroti, and A. Danchin. 1995. Organization of the European *Bacillus subtilis* genome sequencing project. *Microbiology* **141**: 249–255.
- Marshall, E. 1995. A strategy for the sequencing of the genome 5 years early. *Science* **267**: 783–784.
- Ogasawara, N., Y. Fujita, Y. Kobayashi, Y. Sadaie, T. Tanaka, H. Takahashi, K. Yamane, and H. Yoshikawa. 1995. Systematic sequencing of the *Bacillus subtilis* genome: Progress report of the Japanese group. *Microbiology* **141**: 257–259.
- 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.
- Sorokin, A.V., E. Zumstein, V. Azevedo, S.D. Ehrlich, and P. Serror. 1993. The organisation of the *Bacillus subtilis* 168 chromosome region between the *spoVA* and *serA* genetic loci, based on sequence data. *Mol. Microbiol.* **10**(2): 397–405.
- Sorokin, A.V., P. Serror, P. Pujic, V. Azevedo, and S.D. Ehrlich. 1995. The *Bacillus subtilis* chromosome region encoding homologues of the *Escherichia coli mssA* and *rpsA* gene products. *Microbiology* **141**: 311–319.
- Sorokin, A.V., V. Azevedo, E. Zumstein, N. Galleron, S.D. Ehrlich, and P. Serror. 1996. Sequence analysis of the *Bacillus subtilis* chromosome region between the *serA* and *kdg* loci cloned in a yeast artificial chromosome. *Microbiology* (in press).

Received February 8, 1996; accepted in revised form April 8, 1996.