Genomics & Informatics Vol. 1(1) 50-54, September 2003

Genome Mapping of an Extreme Thermophile, *Thermus caldophilus* GK24

Jong Hoon Park[1,3,*1], Byung Chul Park[2], Suk Hoon Koch[2], Joong Soo Kim[2], Jeong Heon Koh[2], Moon Hee Yang[1], Yong Sung Kim[3], Cheorl Ho Kim[3], Myoung Hee Kim[3], Suk Tae Kwon[4], and Dae-Sil Lee[2,3,*2]

[1]Department of Biological Science, Sookmyung Woman's Univ., Seoul, Korea.
[2]Glycobiology Laboratory and [3]Genome Center, Korea Research Institute of Bioscience and Biotechnology, Taejon 305-600, Korea.
[4]Department of Genetic Engineering, SungKyunKwan Univ., Suwon, Korea.

Abstract

Genome of an extreme thermophile, *Thermus caldophilus* GK24 has been analyzed to construct the genomic map. The genomic DNAs encapsulated in agarose gel were digested with *Ssp*I, *Eco*RI, *Spe*I, and *Hpa*I restriction endonucleases, and then the resulting genomic DNA fragments were analyzed by pulsed-field gel electrophoresis. Its restriction map has been constructed by analyzing sizes of the restriction fragments obtained from both complete and partial digestions. The circular form of its genome was composed of about 1.98 Mbp and a megaplasmid. The genomic loci for the genes of xylose isomerase, thioredoxin, tRNA-16S rRNA, 23S rRNA, L5 ribosomal protein, ADP-glucose pyrophosphorylase, DNA-ligase, and *Tca* DNA polymerase were determined by both Southern hybridization and PCR.

Keywords: PFGE; genome map; *Thermus caldophilus* GK24

* Corresponding authors:
[1]E-mail parkjh@sookmyung.ac.kr, Tel +82-2-3452-4718, Fax +82-2-6385-9500
[2]E-mail daesilee@kribb.re.kr, Tel +82-42-860-4110, Fax +82-42-860-4597

Abbreviations: PFGE, pulsed-field gel electrophoresis, PMSF, phenylmethylsulfonyl fluoride

Accepted 13 June 2003

Introduction

The discovery of the genus Thermus opened a detailed study on the physiology of extreme thermophilic bacteria, and provided a unique application of thermo stable enzymes to DNA recombinant technology and industrial uses. *Thermus caldophilus* GK24 strain is a gram-negative, extreme aerophilic and thermophilic bacterium and also can usually grow at 75℃. So far genetic information concerning this strain is limited to only two enzymes, LDH (Koide *et al.*, 1991) and Tca DNA polymerase (Park *et al.*, 1992) reported.

The method of pulsed-field gel electrophoresis (PFGE) of large fragments of chromosomal DNA (Suwanto *et al.*, 1989), generated by using rarely cutting restriction endonucleases, has made it possible to map megabase regions of eukaryotes and whole genomes of prokaryotes (Schwartz *et al.*, 1983 and Bancroft *et al.*, 1989). Approximately 40 bacterial genome maps (Krawiec *et al.*, 1990) have been constructed this way. These maps provide low physical resolution, so their utility has been primarily to demonstrate the mapping strategy and to confirm existing genetic maps (Smith *et al.*, 1987).

In this report, we describe the physical map of *Thermus caldophilus* GK24 genome using pulse-field gel electrophoresis with the restriction fragments. Also several genes of xylose isomerase, thioredoxin, tRNA-16S rRNA, 23S rRNA, L5 ribosomal protein, ADP-glucose pyrophosphorylase, DNA-ligase, and Tca DNA polymerase (Park *et al.*, 1992) were located on the genome map by blot hybridization technique.

Results

Total genome size of the *T.* caldophilus GK24 including extrachromosomal DNAs is about 1.98 Mbp (Table 1), which was estimated, based on the sizes of the restriction fragments of genomic DNA and megaplasmids, which were obtained from their digestions by *Ssp*I, *Eco*RI, *Spe*I, and *Hpa*I endonucleases, respectively. Interestingly it was found that the simple restriction enzymes recognizing A and T nucleotides gave a few restriction fragments, because the genomes of *Thermus* bacteria have high G+C nucleotides. It facilitated the construction of genome mapping. For example, SspI digest of the genomic DNA gave 7 genomic fragments, 699, 574, 393, 309, 150*, 114*, and 64* kb, in which the asterisks indicate the origin of

Table 1. Size of the chromosomal SspI, EcoRI, SpeI, and HpaI restriction fragments from *T. caldophilus* GK24 in kilo base pairs.

SspI		EcoRI		SpeI		HpaI	Ave size
Ss1	699	Ec1	631	Sp1	1,650	Hp1	666
Ss2	574	Ec2	333	Sp2	312	Hp2	335
Ss3	393	Ec3	272	Sp3	159	Hp3	281
Ss4	309	Ec4	225	Sp4	131	Hp4	249
Ss5	150	Ec5	219			Hp5	173
Ss6	114	Ec6	185			Hp6	158
Ss7	64	Ec7	164			Hp7	75
		Ec8	121			Hp8	38
		Ec9	86				
		Ec10	52				
Total :	2,303		2,288		2,252		(1,975)
Genome:	1,975		2,029		1,940	1,975	1,980
Megaplasmid:	328		259		312		300

*1 Shadow box: megaplasmid fragment
*2 ( ): no cutting with *HpaI*

Table 2. Size of the chromosomal partial SspI and EcoRI restriction fragments from *T. caldophilus* GK24 in kilo base pairs.

SspI			EcoRI		
Fragments		Possible combination	Fragments		Possible combination
Ssp1	1065	Ss1+Ss3	Ecp1	819	Ec1+Ec5
Ssp2	965	Ss2+Ss4	Ecp2	754	Ec2+Ec6+Ec4
Ss1	699	Ec1 631			
Ssp3	674	Ss3+Ss4	Ecp3	562	Ec4+ Ec6+Ec7
Ss2	574		Ecp4	487	Ec3+Ec7
Ss3	393		Ecp5	422	Ec4+Ec7
Ss4	309		Ecp6	407	Ec4+6
Ss5	150		Ec2	333	
Ss6	114		Ec3	272	
Ss7	60		Ec4	225	
			Ec5	219	
			Ec6	185	
			Ec7	164	
			Ec8	121	
			Ec9	86	
			Ec10	52	

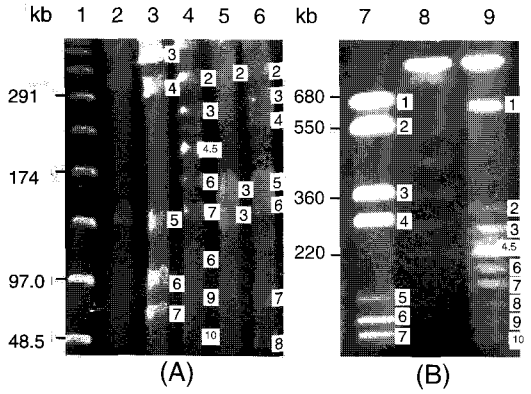

Fig. 1. PFGE of restriction endonuclease digests (*SpeI*, *HpaI*, *EcoRI*, *SspI*) of *T. caldophilus* GK24 genomic DNA. (A) DNA size standard: DNA concatamer (lane 1). Intact genomic DNA of *T. caldophilus* GK24 (lane 2) *T. caldophilus* GK24 genomic DNA digested with *SspI* (lane 3), *EcoRI* (lane 4), *SpeI* (lane 5), *HpaI* (lane 6). The pulse times were 5-25 s for 24 h at 14℃, 200 Volts. (B) DNA size standard: *Saccharomyces cerevisiae* chromosome (lane 8). *T. caldophilus* GK24 genomic DNA digested with *SspI* (lane 7), *EcoRI* (lane 9). The pulse times were 25-75 s for 24 h at 14℃, 200 Volts. The numbers on the right show the positions for the DNA size standard markers of the sizes indicated.

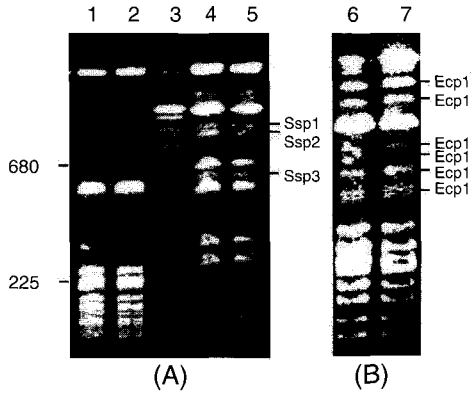

Fig. 2. PFGE of restriction fragments of partial digestion of genomic DNA from *T. caldophilus* GK24 with *SspI* and *EcoRI*. (A) *T. caldophilus* GK24 genomic DNA complete digestion with *SspI* and *EcoRI* (lane 1 and 2). DNA size standard: *Saccharomyces cerevisiae* chromosome (lane 3). *T. caldophilus* GK24 genomic DNA complete digestion with *SspI* (lane 4 and 5) (B) *T. caldophilus* GK24 genomic DNA complete digestion with *EcoRI* (lane 6 and 7). The numbers on the both sides show the positions for the sizes indicated. The pulse times were 25-75 s for 24 h at 14℃, with ramped time, 200 Volts.

megaplasmids. The summation of their sizes was about 1,980 kbp (Table 1 and Fig. 1). Likewise, the digests of *EcoRI* (631, 333, 272, 225, 219, 185, 164, 121*, 86*, and 52* kbp), *SpeI* (1,650, 312*, 159, and 131 kb), and *Hpa* I (666, 335, 281, 249, 173, 158, 75, and 38 kbp) gave about 1,975, 2,029, 1,940, and 1,975 kb respectively. On the other hand, *T. caldophilus* GK24 carries one circular extrachromosommal DNAs, which are shown in PFGE, but we could not detect some of their restriction fragments in PFGE (data not shown). As a result, *T. caldophilus* GK24 has medium size of genome, when compared with genome sizes of other prokaryotes, which range between 600 and 6,000 kbp.

In order to determine the linkage between the restriction

Table 3. Thermus genes used as hybridization probes to PFGE-separated restriction fragments.

Cloned genes	References	Organism	PFGE band			
			SspI	EcoRI	SpeI	Hpa
DNA polymerase	($)	T. caldophilus GK24	Ss1	Ec1	Sp1	Hp3
Thioredoxin	Our lab.*	T. caldophilusGK24	Ss4	Ec1	Sp1	Hp3
DNA-ligase	Dr. Kwon#	T. caldophilus	GK24	Ss1	Ec1	
Xylose isomerase	Our lab.	T. caldophilus	GK24	Ss4		
ADP-glucose pyrophosphorylase	Our lab.	T. caldophilus	GK24	Ss2	Ec1	
tRNAval-16S rRNA	Our lab.	T. caldophilusGK24	Ss3	Ec7		
23S rRNA	Our lab.	T. caldophilus GK24	Ss3	Ec6,Ec7	Sp1	Hp4
Ribosomal protein	Our lab.	T. caldophilus GK24	Ss3		Hp4	

$ Krawiec and Riley, 1990
* Cloned gene fragments in our laboratory
# Received from prof. Kwon S.-T. in SungKyunKwan Univ.

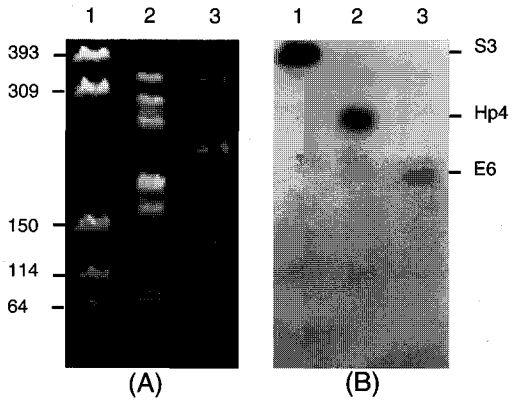

Fig. 3. Physical map of the *T. caldophilus* GK24 genome based on digestion with *Ssp*I, *Eco*RI, *Spe*I, and *Hpa*I. The location of known Thermus genes were determined by hybridization with cloned DNA fragments. (A) Genome (B) MeaplasmidFig. 4. Hybridization analysis with a tRNA-16S rRNA gene probe. T. caldophilus GK24 genomic DNA digested with *Ssp*I (lane 1), *Hpa*I (lane 2), and *Eco*RI (lane 3). (A) PFGE (B) Autoradiography

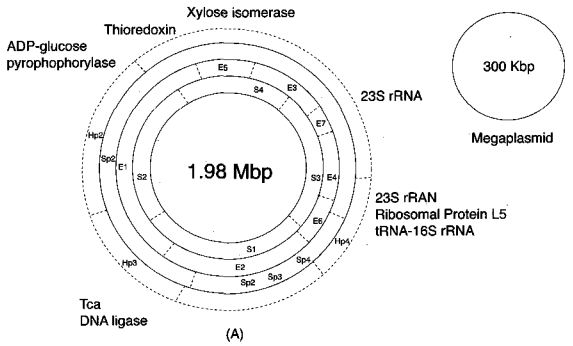

Fig. 4. Hybridization analysis with a tRNA-16S rRNA gene probe. *T. caldophilus* GK24 genomic DNA digested with *Ssp*I (lane 1), *Hpa*I (lane 2), and *Eco*RI (lane 3). (A) PFGE (B) Autoradiography

fragments, partial restriction digests of the genome were analyzed by FPGE. For example, complete and partial digests of the genome with *Ssp*I, gave seven and ten restriction fragments, and complete and partial digests of the genome with *Eco*RI, gave ten and sixteen, respectively (Table 2 and Fig. 2). By analyzing the possible combinations among the complete digests for the partial restriction fragments, a partial digest fragments of 1065 kbp with *Ssp*I endonuclease could be from Ss1+Ss3 (Ssp1); likewise, another fragments of 965 and 674kb are, respectively, from Ss2+Ss4 (Ssp2) and Ss3+Ss4 (Ssp3) (Table 2). Moreover, the double restriction digests could confirm the linkage further. For example, the double digestion with *Ssp*I and *Eco*RI revealed that S2 didn't have *Eco*RI site (Fig. 2). The chromosomal restriction patterns for *Eco*RI involved ten restriction fragments ranging from 631 to 52 kb (Table 1). A partial *Eco*RI digest fragments of T. caldophilus GK24 chromosome was shown in Table 2 and Fig. 2. Furthermore, cross-hybridization method could establish the linkage of the genomic fragments on the physical map. For example, when hybridization with *Ss*1 fragments as a probe was performed to determine the relative position of the *Eco*RI fragments on the chromosome, S1 probe was hybridized on Ec1, Ec2 and Ec6 fragments. As a result, the linkage analysis with *Ssp*I showed the chromosome to have circular form. The complete restriction map of *T. caldophilus* GK24 chromosome with *Ssp*I, *Eco*RI, *Spe*I, and *Hpa*I endonucleases could have been constructed (Fig. 3).

Positions of genes such as carbohydrate-related genes, rRNAs, ribosomal protein, and DNA binding proteins on the restriction map have been determined by hybridization of gene probes and PCR product probes (Table 3). Their positions could be compared with those of microorganisms, and predict their relative positions in the genome of *T. acidophilus* GK24 (Fig. 4).

Discussion

The molecular size of several bacterial genomes such as eubacterium, *E. coli* (Smith, 1987), *thermophilic archaebacterium, Thermococcus celer* (Noil, 1989), *thermophilic Streptococcus thermophilus*, and obligatory aerobic eubacterium, Thermus thermophilus has been reported. Most bacterial genomes comprise one circular chromosome, as determined by genetic mapping and confirmed by physical mapping (Smith *et al.*, 1987). Two possible exceptions to a single circular chromosome draw attention. One is Rhodobacter sphaeroides, which may have two distinct circular chromosomes; the other is the spirochete *Borrelia burgdorferi*, which appears to have a linear chromosome in addition to plasmids, which have covalently closed ends. Generally, chromosome sizes estimated by PFGE and shape determined from ordered libraries of restriction fragments indicated bacterial chromosomes are commonly circular and 1 to 9 Mb. A physical map of the *T. caldophilus* GK24 genome has been constructed by using PFGE and hybridization experiments. The size of the genome, 1.98 Mbp, was smaller than that of *E.coli*. It is close in size to the 2.3-Mbp genome *Hemophilu parainfluenzae* (Kauc and Goodgal, 1989) and *Staphylococuss anguis* (Bourgeois *et al.*, 1989).

The construction of contigs of certain genomic fragment and hybridization experiments can be used to improve the fine genome map and localize some gene cluster of the *T. caldophilus* GK24 genome. Lately we have cloned genes, such as DNA polymerase, Thioredoxin, DNA-ligase, Xylose isomerase, ADP-glucose pyrophosphorylase, tRNAval-16S rRNA, 23S rRNA, Ribosomal protein. Among them, rRNAs are easily useful to be used as probes to determine the abundance, arrangement, composition, and location of rrn loci (Krawiec and Riley 1990). Particularly, *T. caldophilus* GK24 has two rrn loci same as *Thermus thermophilus* (Hartman and Erdman, 1989) and *Pirellula marina* (Liesack and Stackebrandt, 1989), while the members of the Enterobacteriaceae, such as E. coli have seven rrn loci. Use of *EcoR*I restriction endonuclease could find the polymorphism of in the physical map of several *Thermus strains, T. aquaticus* YT-1, *T. thermophilus* HB27, and *T. flavus* AT-62. It would be interesting to compare the physical map of these strains for ladder pattern of *EcoR*I fragments in relation to the evolution of prokaryotes.

Up to date, over 85 microorganisms were sequenced completely. Comparison of a Thermus genome to other prokaryotic genomes should lead to a better understanding of microbial adaptation to extreme conditions, such as hypertemperaure, damaging radiation, and an oxidizing atmosphere. Indeed, the availability of the complete genome sequence for this thermo stable microbe should facilitate a wide range of studies and establish this thermopile as a model organism among the gram-negative bacteria. Also, organization into carbohydrate related gene clusters is an essential and useful of industrial goal. To investigate gene clusters, we are going to focus on the relationship of relative loci about carbohydrate related genes in between *T. caldophilus* GK24 and *E. coli*. Furthermore, we can easily isolated useful genes in industrial goals.

Material and Methods

Strain and culture conditions

Thermus caldophilus GK24 cells (Taguchi *et al.*, 1982) were grown at 75℃, 16 h in medium (pH 7.2) consisting of 0.8% polypeptone, 0.2% yeast extract and a basal salt mixture as described previously (Matsuzawa and Hamaoki, 1983). Then chloramphenicol (180 μg/ml) was added into culture which was maintained for another 4 h.

Preparation of plugs

Agarose plugs containing genomic DNA were prepared as described by Bancroft *et al.* (Bancroft *et al.*, 1989). Cells grown to late log or stationary phase were chilled by swirling in an ice bath and pelleted by centrifugation at 3,500 rpm for 10min at 4℃ in a clinical centrifuge. Cells were then washed by resuspension in 10ml of a buffer (10mM Tris-HCl, 1M NaCl, pH7.6), and followed by centrifugation. After resuspending the cells thoroughly in a suspension buffer (0.01M Tris-HCl, pH 8.0, 0.1M EDTA, 0.02M NaCl), the cells were incubated at 37-42℃, and then were diluted with an equal volume of 1.5% low melting temperature agarose (FMC Bio-Products, Rockland, Maine) in sterile 125mM EDTA solution. The solution was poured into a mould chamber (avoid air bubbles) and cooled the mold at -20℃ for 5min. Then the plugs were transferred to the equal volume of TC lysis solution (6mM Tris-HCl, pH7.6, 1M NaCl, 100mM EDTA, 0.5% Sarkosyl, 1mg/ml lysozyme) and were incubated for 10min at 37℃ with gentle shaking. Discarded the solution and washed the agarose plug three times for 30min in 0.05M EDTA (pH 8.0). And then the plugs were incubated in an equal volume of ESP solution (0.5M EDTA, 1% laural sarcosine, 1mg/ml proteinase K) for 16 h at 50℃ with gentle shaking. In order to remove proteinase K solution completely, the plugs transferred to 1mM TE (pH 8.0) solution containing 1mM PMSF and then incubated for 1h at room temperature. The prepared plugs were stored in 0.05M EDTA (pH 8.0) at 4℃. Restriction enzyme digestion of DNA in agarose plugs. Agarose plugs containing 1mg of genomic DNA were subjected to digestion with restriction endonucleases in 0.1ml of the respective restriction

endonuclease buffer containing 0.01% bovine serum albumin for 20h at 37℃. For these experiments, restriction endonucleases (40 units, ea.), *Ssp*I (B. M.), *EcoR*I (Promega), *Spe*I (B. M.), and *Hpa*I (B. M.) were used. For partial digestion, *Ssp*I (1 U) and *EcoR*I (1 U) were used. After restriction endonuclease digestion, plugs were washed in 50 volume of solution containing TE buffer (pH 8.0). Using a disposable pipette tip, 1/3 of an insert was mounted on the teeth of an electrophoresis comb.

Pulse-field gel electrophoresis (PFGE)

The gel was cast with 1.0% (W/V) Sea-Kem agarose in 0.5 X TBE buffer. The gel was electrophoresesed at 14℃ in a CHEF DRII apparatus (Bio-Rad Laboratories) in field strength of 10V/cm. To resolve restriction fragments over 1000kb, switching times was 120 sec. To separate restriction fragments between 6 and 600 kb, a gradual change of switching intervals from 25 to 75 sec was employed. For separation of fragment sizes between 4 and 200 kb, the gel was run 24 h at 200 V with a ramp of switch time from 5 to 25 sec. After electrophoresis, gels were stained with ethidium bromide for 30 min and photographed with polaroid film.

Hybridization

DNA probes were ^{32}P-labeled using random oligonucleotide primers (Feinberg and volgelstein, 1984). Hybridization experiments with DNA probes were performed at 65℃ in 0.1% SDS (sodium dodecylsulphate), 5 X SSC, 1% laurylsarcosin, and 1% blocking reagent. Gels were washed with 0.1% SDS, 2X SSC (0.3 M NaCl, 30mM sodium citrate, pH 7.0) at 65℃.

Acknowledgments

This work was supported by grant 98-NF-04-01-A-02 from Ministry of Science and Technology of Korea.

References

Bancroft, I., Wolk, C.P. and Oren, E.V. (1989) Physical and genetic maps of the genome of the heterocyct-forming cyanobacterium Anabaena sp. stran PCC 7120. *J. Bacteriol.* 171, 5940-5948.

Baril, C.C., Baranton, R.G., and Saint Girons, I. (1989) Linear chromosome of Borrelia burgdorferi. *Res. Microbiol.* 140, 507-516

Bourgeois, P.L., Muta, M. and Ritzenthaler, P. (1989) Genome comparison oflactococcs strains by pulsed-field gel *electrophoresis. FEMS Microbiol. Lett. 59, 65-70.*

Brock, T.D. (1986) Thermophiles pp. 27-28. John Wiley & Sons, Inc., New York.

Feinberg, A.P. and Vogelstein, B. (1984) "A Technique for radiolabeling DNA restriction endonuclease fragments to high specific activity". *Addendum. Anal. Biochem.* 137, 266

Ferdows, M.S., and Barbour, A.G. (1989) Megabase-sized linear DNA in the bacterium Borrelia burgdorferi, the Lyme disease agent. *Proc. Natl. Acad. Sci. USA*, 86, 5969-5973.

Hartmann, R.K. and Erdman, V.A. (1989) Thermus thermophilus 16S rRNA is transcribed from an isolated transcribed from an isolated transcription unit. *J. Bacteriol.* 171, 2933-2941.

Kauc, L. and Goodgal, S.H. (1989) The size and a physical map of the chromosome of *Haemophhilus parainfuenzae. Gene*, 83, 377-380.

Koide, S., Iwata, S., Matsuzawa, H. and Ohta, T. (1991) Crystallization of allosteric L-lactate dehydrogenase from Thermus caldophiluous and preliminary crystallographic data. *J. Biochem.* 109, 6-7.

Krawiec, S. and Riley, M. (1990) Organization of the bacterial chromosome. *Microbiol. Rev.* 54, 502-539.

Liesack, W. and Stackebrandt, E. (1989) Evidence for unlinked rrn operons in the planctomycete pirellula marina. *J. Bacteriol.* 171, 5025-5030.

Matsuzawa, H., Hamaoki, M. and Otha, T. (1983) *Agri. Biol. Chem.* 47, 25-28.

Noil, K.M. (1989) Chromosome map of the thermophilic archaebacterium Thermococcus celer. *J. Bacteriol.* 171, 6720-6725.

Park, J.H., Kim, J.S., Kwon, S.-T. and Lee, D.-S. (1992) Purification and characterization of Thermus caldophilus GK24 DNA polymerase. *Eur. J. Biochem.* 214, 135-140.

Schwartz, D.C., Saffran, W., Welsh, J., Haas, R., Goldenberg, M. *and Cantor, C.R. (1983) New techniques for purifying large* DNAs and studying their properties and packing. *Cold Spring Harbor. Symp. Quant. Biol.* 47, 189-195.

Smith, C.L., Econome, J., Schutt, A., Klco, S. and Cantor, C.R. (1987) A physical map of the E. coli K-12 genome. *Science* 236, 1448-1453.

Suwanto, A. and Kaplan, S. (1989) Physical and genetic mapping of the Rhodobacter sphaeroides 2.4.1. genome :genome size, fragment identificaion, and gene localization. *J. Bacteriol.* 171, 5840-5849.

Taguchi, H., Yanashita, M., Matsuzawa, H. and Otha, T. (1982) Heat-stable and fructose 1,6-bisphosphate-activated L-lactate dehydrogenase from an extremely thermophilic bacterium. *J. Biochem.* 91, 1343-1348.