

특집: 미생물계통 및 분석방법(II)

*Bacillus subtilis*의 genome 연구현황 및 전망

박 승 환

KIST생명공학연구소 응용미생물연구부

*Bacillus subtilis*는 *Bacillus*속 균주의 type species로서 *Escherichia coli*와 함께 원시핵세포생물 연구를 위한 주요 모델로서 이용되어 왔다. 많은 연구자들이 *B. subtilis*에 대해 관심을 가지게 된 동기는 포자형성이라고 하는 비교적 단순한 developmental system을 가지고 있고 산업적으로 중요한 효소와 항생물질 등을 다량 생산하고 분비한다는 점 때문이다. 포자를 형성하는 과정에서 하나의 세포가 비대칭의 두 구역으로 나누어지면서 어떻게 공간적 시간적으로 차별적인 유전자발현이 일어나며 또 forespore와 모세포사이에 communication이 일어나는지가 주된 관심사였다. 산업적 이용면에서 *Bacillus*속 균주들이 생산하는 세가지 중요한 산물로서 효소, 항생제 및 살충제를 들 수 있다. Proteases, amylases, glucanases 및 cellulases 등의 효소가 다양한 *Bacillus*균주들에 의해 생산되는데 이들에 의해 생산된 효소제품은 세계 산업용 효소시장의 3분의 2 이상을 점유하고 있다. 고생산성균주의 개발 및 발효기술의 향상과 더불어 *B. subtilis*균이 GRAS(generally regarded as safe) 미생물로 인정됨에 따라 *Bacillus*균의 중요성은 더욱 증대되었으며 외래 유용단백질, 항대사능 및 약리활성이 있는 다양한 2차 대사산물 생산숙주로서의 개발에 더욱 박차를 가하게 되었다. 산업적으로 중요한 세번째 부류의 산물은 *B. thuringiensis*, *B. sphaericus* 및 *B. popilliae* 등의 균주가 생산하는 살충성 독소단백질이다. 이들은 대개 포자형성이 진행되는 동안 parasporal inclusion 형태로 생산되는데 나비목, 파리목 및 딱정벌레목 등의 해충에 특이적으로 살충성을 나타낸다(2,3).

*B. subtilis*의 전체 genome sequencing의 중요성

어떤 미생물의 전체 genome sequence가 밝혀졌을 때 그 생명체에 대한 연구 방식은 크게 달라질 것이며 연구진행도 가속될 것이다. 그러나 현재의 기술수준과 기타의 여건상 현재 연구되고 있는 모든 미생물을 대상으로 genome sequencing을 수행할 수는 없으며 순수과학적 측면에서나 응용적 측면에서 얻을 수 있는 이익이 큰 대상들을 선정할 필요가 있다. 이러한 관점에서 선정된 대표적인 것이 *E. coli*(14)와 *B. subtilis*이다.

*B. subtilis*가 갖는 세가지 장점은 첫째, Gram 양성세균의 연구모델로서 중요한 위치를 차지하고 있다는 점, 둘째, 산업적으로 유용성이 큰 균주라는 점과 더불어 셋째, 전체 genome 크기가 4188 kb(8)로 작다는 점이다. *B. subtilis*는 유전자 조각이 비교적 용이한 까닭에 replication, transcription, translation, genetic recombination, chemotaxis, 포자형성 등 다양한 생물학적 현상들에 대한 연구가 활발히 진행되어 왔다. 연구자들은 이러한 현상들이 일어나는 기작 뿐만 아니라 이에 관련된 유전자들을 밝히고자 노력하고 있는데 염색체상에 존재하는 모든 유전자에 대한 자세한 정보가 마련되면 이러한 연구에 커다란 도움이 될 것이다. 여러 *Bacillus* species가 산업적으로 중요하게 인식되고 있는데 반해 *B. subtilis*만큼 쉽게 다룰 수 있는 체제가 갖추어지고 고전적 방법 또는 분자유전학적으로 연구가 된 균주는 없다. 따라서 *B. subtilis*는 *Bacillus*속 미생물연구의 모델로서의 의미가 크며 이 균주에 대한 자세한 정보 및 연구 성과는 산업적으로 중요하면서 연구가 부진한 다른 *Bacillus*균의 연구에 도움이 될 것이다. *B. subtilis*와 *E. coli*는 적어도 20억년 전에 서로 다른 두 종으로 갈라지게 된 것으로 추정되는데(16) 현재는 현격히 다른 두 환경, 토양과 장내를 각자의 서식처로 삼고 있다. 이 두균주의 gene content와 genome 구성을 비교함으로써 두균주에 대해 현재까지 밝혀진 유사점과 상이점을 보다더 깊이 이해할 수 있게 될 것인데 이 또한 genome project를 통해 얻을 수 있는 커다란 수확이 될 것이다. 이밖에 *B. subtilis*에 대하여 연구를 수행하는 연구팀이 매우 많고, genome의 G+C 함량이 40%로 낮으며 repetitive DNA가 적고, 자세한 genetic map 및 physical map이 구비되어 있으며, 이미 여러 지점의 sequence가 알려져있는 점 등이 본 project의 형성을 촉진한 요인이 되었다.

B. subtilis genome project의 형성 및 재원

*B. subtilis*에 대한 genome sequencing project는 당초 미국의 5개 연구팀과 유럽의 5개 연구팀이 연합하여 추진하려고 했던 것인데 연구비 확보가 유보된 미국팀들이 제외된채 1989년 9월 유럽의 5개 연구팀들(A. Danchin & G. Rapoport, J. Errington, A. Galizzi & A. Albertini, S. D. Ehrlich, K. Devine

팀)에 의해 시작되었다. 이들은 컨소시엄을 형성하여 European Commission(EC)의 SCIENCE 프로그램 지원하에 physical map 작성, DNA libraries 제작 및 pilot scale의 systematic sequencing을 1차적인 목표로하여 연구를 시작하였고 당시 Coordinator로는 Pasteur 연구소의 R. Dedonder가 맡았다(10). 1993년 4월 4개연구팀(S. Bron, C. Harwood & P. Emmerson, R. Mellado, G. Grandi팀)이 EC의 BIOTECHNOLOGY 프로그램의 지원으로 합류하였고, 스위스의 D. Karamata 연구팀이 스위스 정부의 지원을 받아 European Union member가 아닌 최초의 연구팀으로서 이 project에 합류하였다. 그후 EC의 지원하에 Europe의 7개팀(J. Haiech & F. Denizot, B. Oudega, I. Connerton, R. Borriss, B. Joris, S. Seror, K. Entian 팀)이 추가로 합류하게 되었고 핀란드-미국 합자회사인 Genencor International사(San Francisco, CA, USA)와 Novo-Nordisk Biotech 사(Davis, CA, USA)가 자체의 연구비로 참여하게 되었다. 한편, Japanese Human Genome Project의 지원하에 1991년부터 *B. subtilis*의 genome sequencing을 시작한 일본의 7개 연구팀이 컨소시엄을 형성하여 EU의 연구팀들과 긴밀한 협조체제하에 일을 분담하고 있으며 본 생명공학연구소의 1개 연구팀이 Pasteur연구소의 Dr. A. Danchin 및 Dr. F. Kunst와의 공동연구형태로 1995년 5월부터 본 과제에 참여하게 되었다. 이렇게하여 현재 총 27개의 연구팀이 참여하는 international project로 발전하였다(6,9). 현재 유럽 컨소시엄의 coordinator는 Dr F. Kunst가 맡고 있으며 Dr R. Dedonder와 Dr C. Anagnostopolus가 전체 project의 scientific advisor로서 참여하고 있다.

염색체 DNA구역의 할당

모든 참여팀들은 Dr Anagnostopoulos가 제공한 *B. subtilis* strain 168의 염색체 DNA를 sequencing하도록 되어 있다. *B. subtilis*에 대한 상세한 genetic map과 SfiI과 NotI을 이용한 physical map이 조기에 완성되고(1,8) mapping된 유전자중 많은 수의 염기서열이 이미 밝혀짐에 따라 여러 연구팀에게 정확한 DNA region을 할당할 수 있었고 나아가 cloning 과 sequencing을 수행하는데 있어서 유용하게 활용되어왔다. 각 팀당 10~15Å(대략 100~150 kb)의 염색체부위가 할당되었으며(그림 1) 연구팀당 일년에 20 kb 이상의 새로운sequence data를 제출하도록 되어있다.

Cloning and sequencing

각 연구팀들은 할당된 부위의 염색체DNA를 각자 확보하는 것을 원칙으로 하고있다. 여러 연구팀들은 그동안의 경험을 통해 sequencing project의 bottleneck이 sequencing 단계가 아니라 cloning과 data 해석 이라는 사실을 알게 되었으며 염색체DNA 및 이에대한 일련의 sublibrary를 얻기 위해 다음과 같은 전략을 사용하고있다.

- (a) plasmid, λ 및 YAC vector 를 이용한 cloning,
- (b) YAC recombinant vector를 이용한 λ clones 의 odering,
- (c) integrative plasmids와 marker rescue를 이용한 genome walking,
- (d) PCR에 의한 *in vitro* amplification.

이렇게 여러 방법을 함께 사용하게 된 이유중 하나는 *E. coli*에 cloning된 *B. subtilis* DNA가 불안정하거나 cloning이

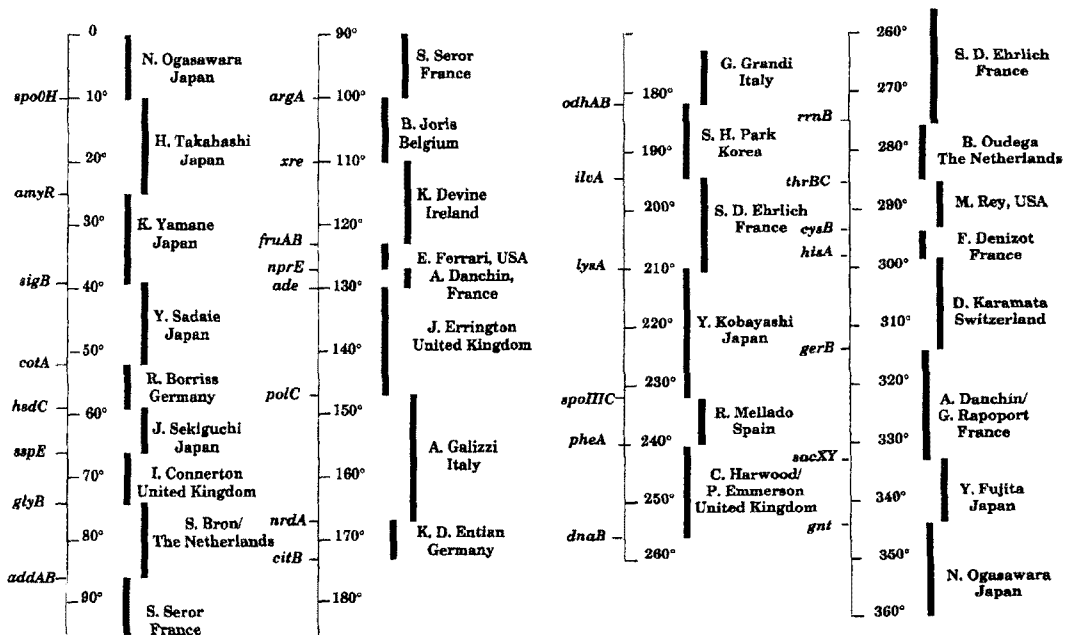


Fig. 1. The international project of sequencing the *Bacillus subtilis* genome.

되지 않는 경우가 많기 때문이다. 몇몇 연구팀들은 λ 나 cosmid vector를 이용한 완전한 library를 만들고자 하였으나 실패한 경험이 있다. Plasmid rescue방법에 의해 cloning을 할 경우에 이들은 ColE1 plasmids의 copy수를 낮게 유지시키는 TP611을 주로 사용하였다(7). 여러가지 방법을 사용해서도 *E. coli*에 cloning이 되지 않는 DNA fragments의 경우에는 marker rescue 방법을 이용하여 *B. subtilis*를 숙주로 하여 cloning 하였다. YAC vector를 이용한 library가 가장 성공적인데 S.D. Ehrlich 등에 의한 library의 경우 *B. subtilis* 전체 chromosome의 98%를 포함하는 것으로 보고되었다(4).

Sequencing 또한 각 팀들이 나름대로의 방법을 사용하여 수행하고 있는데 대부분은 sonication이나 DNase I을 이용, genomic clones으로부터 random subfragments를 만들어 sequencing하는 방식을 택하고있는 반면 몇몇 연구팀은 systematic primer walking 방식을 택하고 있다. Sequencing은 대부분 자동화된 sequencing machine을 이용하여 수행하는 추세이다.

Data handling

각 연구팀들이 얻은 sequence data는 일단 불란서와 일본에 설치된 sequence 저장소에 보내도록 되어있다. 여기서 다른 연구팀에 의해 15% 정도의 sequencing을 반복하여 확인과정을 거친다음 정식data로 채택되며 9개월간의 유보기간을 거쳐 GenBank나 EMBL 등의 공공의 database와 *B. subtilis* 전용 database인 SubtiList에 수록된다(11).

성 과

대부분의 유럽 연구팀들은 매년 20 kb 이상을 sequencing하고 있으며 이미 확보한 contig에는 127 kb, 105 kb, 60 kb, 50 kb와 42 kb의 크기를 갖는 것들이 있다. 현재 평균 error rate는 2 kb당 1 bp이다. 1995년초까지 유럽 연구팀들이 piecemeal approach나 systematic approach에 의해 밝혀낸 DNA sequence는 각각 600 kb와 500 kb이며 1996년말까지는 1.5 Mb를 달성할 것으로 내다보고 있다(9). Japanese project를 살펴보면 7개 연구팀이 총 1.3 Mb의 부위를 할당받아 수행하고 있는데 1995년초까지 700 kb에 달하는 clone을 확보하였고 약 500 kb의 sequencing 결과를 확보한 것으로 보고하였다. 이들은 주로 λ phage vector를 이용하여 cloning을 하였고 문제가 있을 경우 inverse PCR method를 사용하여 해결하고 있는데 1997년까지 할당받은 전체부위를 완료할 수 있을 것으로 보고 있다(12). 본 연구팀은 현재 약 40 kb의 염색체 DNA를 확보하여 이중 8 kb에 대한 sequencing을 완료하였으며 앞으로 sequencing과 더불어 genome의 function에 대한 연구도 병행해나갈 것이다.

생물산업

전망 및 향후과제

본 과제는 1997년말까지 전체 genome sequencing을 완료할 계획으로 추진되고 있으며 현재의 추진속도로 볼때 가능할 것으로 생각된다. 같은 부위의 sequencing을 완료한 연구팀들은 본 project의 조기 완료를 위해 진도가 부진한 다른 부위를 추가로 할당받아 일을 진행하고 있다.

지금까지의 systematic genome sequencing을 통해 밝혀진 가장 중요한 사실은 새로 드러난 open reading frame의 30~50%가 기존의 databases와 상동성을 가지지 않으며 현재 그 기능이 규명되지 않은 상태라는 것이다(7,12,15,5,13). 이들은 고전적인 유전학적 연구방법으로는 밝혀내기 어려운 생명유지에 필수적인 유전자이거나, 복수의 조절기작을 갖는 생합성 경로의 한 member라서 변이가 일어나도 별로 영향이 없는 유전자이거나, 아직까지 알려지지 않은 새로운 biological process에 관여된 유전자일 수 있으며 앞으로 많은 연구가 이루어져야할 대상들이다. 특히 이들 유전자의 기능을 알아내는 것이 큰 과제인데 전체 genome sequence가 밝혀지면 보다 체계적인 접근이 가능해질 것이며 여러 돌연변이주의 유전자 발현을 쉽게 조사할 수 있는 자동화된 분석시스템개발이 요청된다. 기능적으로 중요한 유전자들은 일반적으로 잘 보존되는 것으로 알려지고 있다. 따라서 세균간의 gene content와 genome구성을 비교 분석함으로써 이러한 잘 보존되어온 유전자들을 자세히 살펴볼 수 있게 될 것이다. 또한 어떤 세균의 새로운 유전자가 기능이 잘 밝혀진 다른 세균의 유전자와 상동성이 있을 경우 쉽게 그 기능을 분석해 들어갈 수 있을 것이다. 결국은 컴퓨터분석을 통해 세균들의 핵심유전자와 특수한 환경하에서만 필요한 유전자들을 구분할 수 있을 것이며 따라서 좀더 풍부한 비교분석자료를 확보하기 위해 *E. coli*와 *B. subtilis* 이외에 다양한 세균들에 대한 genome sequencing이 계속되어야 할 필요가 있다.

참고문헌

1. Anagnostopoulos, C., Piggot, P. J. and Hoch, J. A. 1993. The genetic map of *Bacillus subtilis*. In *Bacillus subtilis and Other Gram-Positive Bacteria: Biochemistry, Physiology, and Molecular genetics*, pp. 425-461. Sonenshein, A. L., Hoch, J. A., and Losick, R.(eds). Washington, D. C., American Society for Microbiology.
2. Aronson, A. I. 1993. Insecticidal toxins. In *Bacillus subtilis and Other Gram-Positive Bacteria: Biochemistry, Physiology, and Molecular genetics*, pp. 953-964. Sonenshein, A. L., Hoch, J. A. and Losick, R.(eds). Washington, D. C., American Society for Microbiology.
3. Aronson, A. I., W. Beckman and P. Dunn. 1986. *Bacillus thuringiensis* and related insect pathogens. *Microbiol. Rev.*

- 50, 1-24.
4. Azevedo, V., Alvares, E., Zumstein, E., Damiani, G., Sgaramella, V., Ehrlich, S.D. & Seror, P. 1993. An ordered collection of *Bacillus subtilis* DNA segments cloned in yeast artificial chromosomes. *Proc. Natl. Acad. Sci. USA* **90**, 6047-6051.
 5. Cosmina, P., F. Rodriguez, F. de Ferra, G. Grandi, M. Pereggo, G. Venema and D. van Sinderen. 1993. Sequence and analysis of the genetic locus responsible for surfactin synthesis in *Bacillus subtilis*. *Mol. Microbiol.* **8**, 821-831.
 6. Devine, K. M. 1995. The *Bacillus subtilis* genome project: aims and progress. *Trends in Biotechnology*, **13**, 210-216.
 7. Glaser, P. and others. 1993. *Bacillus subtilis* genome project: cloning and sequencing of the 97 kb region from 325 Å to 333 Å. *Mol Microbiol.*, **10**, 371-384.
 8. Itaya, M. 1993. Physical Map of the *Bacillus subtilis* 168 chromosome. In *Bacillus subtilis* and Other Gram-Positive Bacteria: Biochemistry, Physiology, and Molecular genetics, pp. 463-471. Sonenshein, A. L., Hoch, J. A. and Losick, R. (eds). Washington, D.C., American Society for Microbiology.
 9. Kunst, F., A. Vassarotti and A. Danchin. 1995. Organization of the European *Bacillus subtilis* genome sequencing project. *Microbiology*, **141**, 249-255.
 10. Kunst, F. and K. Devine. 1991. The project of sequencing the entire *Bacillus subtilis* genome. *Res. Microbiol.* **142**, 905-912.
 11. Moszer, I. Glaser, P. and Danchin, A. 1995. SubtiList: a relational database for the *Bacillus subtilis* genome. *Microbiology* **141**, 261-268.
 12. Ogasawara, N. and others. 1995. Systematic sequencing of the *Bacillus subtilis* genome: progress report of the Japanese group. *Microbiology* **141**, 257-259.
 13. Ogawa, K., E. Akagawa, K. Nakamura and K. Yamane. 1995. Determination of a 21548 bp nucleotide sequence around the 24° region of the *Bacillus subtilis* chromosome. *Microbiology* **141**, 269-275.
 14. Sofia, H. J., Burland, V., Daniels, D. L., Plunkett, G., III & Blattner, F. R. 1994. Analysis of the *Escherichia coli* genome. V. DNA sequence of the region from 76 to 81.5 minutes. *Nucleic Acids Res.* **22**, 2576-2586.
 15. Sorokin, A., E. Zumstein, V. Azevedo, S.D. Ehrlich and P. Serror. 1993. The organization of the *Bacillus subtilis* 168 chromosome region between the *spoVA* and *serA* genetic loci, based on sequence data. *Mol. Microbiol.* **10**, 385-395.
 16. Woese, C. R. 1987. Bacterial evolution. *Microbial Rev.* **51**, 221-271.

제일제당 종합연구소 소식지 창간

제일제당 종합연구소(소장 백운호 부사장)는 기업의 자체 연구개발에 대한 최신 정보들을 대 내외에 알려 산학연 협력체제를 촉진하고자 하는 의미에서 2월말 연구소 소식지를 창간하였다. "무한탐구"라고 명명된 이 소식지는 계간으로 발간될 예정이며, 5×7 변형판 총 16면으로 구성되어 전문 상식, 유명 연구소 소개, 신기술 및 신제품, 논단, 과학 산책 등의 내용으로 이루어져 있다. 회사내 관련 부서를 비롯하여 당 연구소와 관련이 있는 국내 연구기관 및 대학 등에 배포할 예정이다. 정기적으로 당 연구소 소식지를 받아보기를 희망하시는 분은 제일제당 종합연구소 길광훈 박사(Tel: 0336-39-4322, Fax: 0336-32-2784)로 연락바랍니다.