

Micro Column ①

NGS시대와 미생물 연구의 재조명

정해영

한국생명공학연구원 바이오합성연구센터 책임연구원

지금으로부터 약 20년 전 대학원생 시절, 학교 신문에 연구실 소개의 글을 실은 적이 있다. 생명과학 전공의 여느 대학원 연구실과 마찬가지로 아침을 맞으면서 램이 부산하게 움직이기 시작하는 모습을 나름대로 생동감 있게 묘사한 것으로 기억한다. Autoclave

는 김을 뿜으며 쉴 새 없이 배지와 초자 및 일회용 기구를 토해 내고, 약간 요령을 피울 줄 아는 선배들은 바지런한 후배들에게 “LB plate 하나만 주라” 또는 “Buffer 만들어 놓은 것 없냐?” 등등 자질구레한 부탁을 하곤 했었다. 그 중에서 수혜자의 수고를 가장 크게 덜어 주는 일은 바로 DNA sequencing 샘플을 로딩하기 위한 여분의 lane을 얻어내는 것이었다.

요즘에는 DNA sequencing 샘플을 근처의 회사에 보내기만 하면 하루나 이틀 후 염기서열 정보로 깨끗이 전환된 결과물을 받을 수 있다. 이제는 연구 및 실험의 상당 부분을 외부 회사에 아웃소싱하는 것이 일반화 되었고, 생명공학이 추구하는 진정한 산업화와는 아직 약간의 거리가 있지만 대학과 연구소에서 직접 진행하는 것 외에는 대안이 없었던 실무

의 많은 부분이 ‘제품화’ 되어 나라의 경제를 돌아가게 하고 고객, 즉 일선 연구자들은 좀 더 가치 있는 일에 몰두할 수 있게 된 것은 매우 바람직한 변화라 아니할 수 없다.

당시 저자가 몸담았던 연구실은 폴리아크릴아미드 slab gel을 이용한 DNA sequencing 기술에서는 국내 최고 수준이라 해도 과언이 아닐 정도로 깨끗한 결과를 뽑아내고 있었다. 노출 및 현상이 완벽하게 되어 선명하게 검은색 띠가 줄줄이 배열된 X-ray 필름을 들고 자랑스럽게 암실을 나서는 것은 매우 즐거운 일이었다. 암실 서랍 속에서는 아무것도 노광이 되지 않아 투명하게 현상되어 버린 망친 필름(이를 흔히들 ‘책받침’ 이라고 하였다)이나, 숨은 사연을 알 수 없으나 쓰지 못하게 된 뿌연 생필름이 몇 장씩 발견되고는 해서 암실의 어둠 속에서 머리를 싸매고 좌절하고 있을 어느 대학원생의 고뇌를 미루어 짐작하게 하는 일이 흔히 벌어지고는 하였다.

USB(컴퓨터와 외부 기기를 접속하는 Universal Serial Bus가 아니라 United States Biochemical로서 2008년 Affymetrix에 인수됨)의 Sequenase 키트는 정말 아껴 써야만 하는 보물 같은 존재였고, 수요를 잘못 맞추어 Amersham사의 방사성 동위원소가 너무 많이 남게 되면 연구실 재정을 담당하는 선배로부터 핀잔을 듣기도 했었다. 향은 수조 앞에서 타이머를 켜 놓고 template DNA와 primer를 섞은

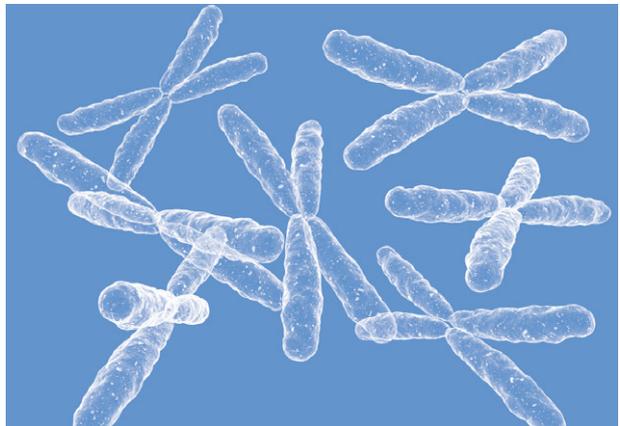
annealing mixture를 준비하여 labeling 반응을 시킨 뒤 이를 termination mixture를 이미 분주한 4개의 tube에 나누어 담는 모습은 후에 시퀀싱 반응물을 분리하기 위해 두꺼운 두 장의 유리판 사이에 기포가 생기지 않도록 폴리아크릴아미드 용액을 조심스레 붓고, 굳은 뒤 이를 전기영동 탱크에 세워서 셋업을 하기까지의 긴장된 모습과 견줄 정도로 진지하고도 성스러운 과정이었다. 전기영동을 하는 수 시간 동안 역시 안심할 수 있는 시간이 아니었다. 당시는 아크릴로 직접 제작한 전기영동 탱크가 전문 회사에서 만들어진 좋은 품질의 탱크로 막 교체가 되던 시절이었다. 아크릴을 잘라서 클로로포름을 접착제로 하여 붙여 만든 전기영동 탱크는 새기 일쑤였고, 종로 금은방에서 주문해 온 백금선은 종종 끊어지기도 하였다. 6시간에서 길게는 10시간 가까이 전기영동을 마친 뒤에는 두 장의 유리판을 조심스레 분리하여 종이장만큼 얇은 gel을 3MM paper에 발라내야 하는 고난이도의 작업이 기다리고 있었고, 여기까지 무사히 진행을 했다 하더라도 진공펌프에 걸어서 gel을 말리는 과정 역시 많은 운이 따라야 했다. 완전히 건조되기 전에 진공펌프를 잘못 건드려 진공이 빠지더라도 하면 gel은 가뭄에 드러난 강바닥처럼 순식간에 갈라 터지면서 하루 동안의 모든 노력을 허사로 만들어 버렸으니까 말이다.

90년대 중반이 되면서 인간 유전체 프로젝트 (http://www.ornl.gov/sci/techresources/Human_Genome/home.shtml)의 성과가 하나 둘씩 가시화되었고, 이를 가능하게 하는 전후방 산업이 성장하기 시작하였다. 졸업 후 자리 잡은 첫 직장에서도 5대의 ABI 3700을 동료들과 같이 셋업하고 미생물의 whole-genome shotgun sequencing을 위한 작업 프로세스를 구축하면서 한때는 국내 최대 규모의 유전체 정보 생산 설비를 갖추었노라고 자부심을 갖기도 하였으나 이는 그다지 오래 가지 않았다.

2005년 454/Pyrosequencing 기술이 상용화되면서 세상은 정말 하루가 다르게 변하기 시작하였다(1). 454/Pyrosequencing 뒤에는 Solexa(Illumina), 그리고 Ion Torrent PGM... NGS data는 기존의 유전체 염기서열 해독

뿐 아니라 transcriptome, DNA-protein interaction, 생명체 간 상호 작용 등 생명과학 분야의 거의 전 분야에 걸쳐서 기존의 실험 기법을 점차 대체해 나가고 있다. 이와 더불어서 open source software의 확산으로 NGS data의 분석에 활용할 수 있는 소프트웨어도 더욱 많아지고 있다. 미생물 유전체의 de novo assembly 분야만 보더라도 Sanger chemistry 시절에는 phrap이나 cap3 정도만 능숙하게 다룰 줄 알면 거의 대부분의 용도를 충족할 수 있었지만, NGS data의 경우에는 SOAPdenovo, CLC Genomics Workbench, Velvet, Mira, ABySS 등 다양한 프로그램을 한번 씩은 다 만져 보고 그 결과를 비교할 정도는 되어야 하는 시대가 되었다(2, 3).

염기서열 해독에 소요되는 비용이 획기적으로 절감되고, 예전에는 상상하지도 못할 저렴한 비용에 수백 배수 이상의 미생물 유전체 서열 데이터를 해독하게 된 요즘, 연구자들은 더 많은 자유? 데이터에서 가치 있는 지식을 뽑아내기 위해 고민하고 사색할 수 있는 자유와 여유? 를 누리게 되었을까? 우리는 아직 이러한 철학적인 질문에 대답할 수준에는 이르지 못한 것 같다. 단지 과학기술자로서, 지금껏 객관적이고도 정량적인 질문에 대해서만 일정 수준의 유의 수준을 갖는 예견을 할 수 있을 뿐이다. 예를 들자면 우리는 “현재의 Wi-Fi 환경에서 이론적으로 몇 MB/s의 전송이 가능한가?” 또는 “Illumina mate-pair library 제작 방법으로 최대 몇 kb에 해당하는 insert를 안정적으로 수용할 수 있는가?” 라는 질문에 대해 매우 구체적으로 대답할 수는 있지만, “우리 자



신과 배우자, 그리고 자녀의 유전체 염기서열을 미리 아는 것이 우리를 진정 행복하게 할 수 있는가?” 에는 아직 대답하기 어려운 것처럼.

이제는 차세대 염기서열 해독 기술을 더욱 발전시켜서 단일 DNA 분자에서 일어나는 중합 반응을 직접 검출하여 염기서열 정보로 바꾸어 주는 제3세대 염기서열 해독 기술이 실현되고 있다. 몇 년 전부터 연구자들의 가슴을 설레게 했던 Pacific Biosciences의 RS는 아직 시장에서 큰 수익을 내고 있지 못하는 것 같은데, 이번에는 단일 가닥 DNA가 미세한 구멍을 통과하면서 나타나는 전기 전도도의 차이를 염기서열 정보로 변환해 내는 장비가 나왔더니 정말 놀라울 따름이다(4). 사실 제2세대 장비조차 어느 한 가지 platform으로 확립되지도 않은 상태에서 다음 세대의 기술이 시장에 선보일 준비를 하고 있다는 것은, 실수요자 측면에서는 대단히 곤혹스러운 일이기도 하다. 도전적인 연구자에게는 다양한 기법을 체험해 보면서 최적의 답을 찾아나가는 즐거운 과정일

수도 있지만, 대부분의 연구자들에게는 200년 전 산업혁명시대를 저항하였던 노동자들의 러다이트 운동을 떠올릴 수도 있을 것이다. 물론 차세대 염기서열 해독 장치가 지금 당장 우리의 일자리를 위협하는 것은 아니고, 단지 하루가 다르게 쏟아져 나오는 최신 기법에 익숙해지기 위해 항상 준비하는 모습이어야 한다는 현실이 때로는 부담스럽기도 하다.

디지털·모바일 기기가 인터넷이라는 새로운 세상을 만나면서 누구나 정보를 생산하고 공유할 수 있는 세상이 되었다. 프로와 아마추어의 경계는 점차 모호해지고 있으며, 쓰레기 정보가 양산되고 있다는 우려 속에서도 일정 규모 이상으로 모인 정보는 스스로에게 의미를 부여해 가면 때로는 세상을 바람직한 방향으로 이끌어 나가기도 한다. 유전체 정보 생산도 마찬가지로 모습을 하고 있다. 과거에는 사진 촬영이 전문가의 영역이었지만, 디지털 카메라가 대중화된 지금은 누구나 영상을 생산하고, 공유하고, 즐기고 있으며 또 그중에는 꽤 수준이 높은 사진도 많다. 마찬가지로 전문 시퀀싱 센터를 직접 운영하거나 큰 규모의 연구비를 확보한 연구실이 아니라면 불가능했던 유전체 해독 연구를 이전 어떤 규모의 랩에서라도 할 수 있는 일이 되었다. 우리 연구팀만 하여도 10개에서 20개 가량의 미생물 유전체를 모아서 대략 분기마다 한번씩 시퀀싱 작업을 보내고 있다. 약 한 달 뒤에 결과가 나오면, 필자를 비롯하여 몇 명의 연구원은 수십 톤의 원광석을 뒤져서 몇 조각의 금을 캐듯이 힘을 들이지만 보람 있는 일에 또 다시 매달리게 될 것이다. 이렇게 읽어 내는 미생물 유전체가 지금 얼마짜리의 경제적·학술적 가치가 있는지는 당장 판별하기 어렵다. 그러나 조금해 할 것은 없다. 당장은 무질서하고 혼란스러워 보인다 해도, 우리는 아직 지식 발굴의 시대를 살아가야 한다고 믿는다.



그림 1. 그림 . 2012년 6월 14일자 Nature 표지. 이번 표지 디자인은 Scientific artist인 Joana Ricou의 원작 'Our Self-Portrait: the Human Microbiome'에서 영감을 받아 작성되었다. Broad Institute가 공개한 보도자료는 <http://www.broadinstitute.org/news/4199>를 참조할 것.

지난 6월 세계 주, Human Microbiome Project (HMP) Consortium의 종합적인 연구 성과 14편이 Nature와 Genome Biology, 그리고 PLoS 등의 저널에 소개되었다. 장과 피부 등 인체의 주요 기관에 서식하면서 인체와 밀접한 관계에 있는 미생물집단의 연구는 유전체학 연구의 기술 혁신이 아니면 불가능한 것이었다. 인간 유전체 프로젝트와 같

은 Big Science의 이면에는 항상 투입 대비 효용을 부르짖는 반대론자가 있었고, HMP도 마찬가지였을 것이다. 242명의 건강한 성인으로부터 채취한 미생물 시료의 다양성에서 어떠한 실용적인 결과가 파생될지는 아직 알기 어렵다. 그러나 한 가지 확실한 것은, 인간 유전체 프로젝트와 기존의 생명과학을 통해 우리 인간이 갖춘 분자적·유전적 기초를 알아 왔다고 한다면, 이제는 **superorganism**으로서의 인간을 구성하는 또 다른 가족인 미생물에 관심을 가질 때가 된 것이다. 아직은 연계성을 알기 어려운 대용량의 데이터들이지만, 이들이 모이고 연구자 커뮤니티에 의해 끊임없이 반복되면서 지금으로서는 가늠하기 어려운 ‘지혜’가 도출되는 데에는 많은 시간이 걸리지 않을 것이다.

생명공학의 응용을 위한 도구로서만 인식되던 미생물 연구가, NGS와 HMP 시대를 맞이하여 인간과의 상호작용이란 측면에서 새롭게 조망되는 것은 미생물 유전체를 다루는 일선 연구자로서 매우 반가운 일이 아닐 수 없다. 미생물이 우리에게 주는 교훈은 다양성과 적응성이 아닐까? NGS 시대의 한 복판에 서서, 하루가 다르게 출시되는 새로운 장비와 쏟아지는 데이터 터미에 피로감이 밀려온다면 이렇게 되뇌어 보자. 다소 혼란스럽더라도 그대로 굴러가게 놔 두어 보자(5). 그 속에서 새로운 질서와 가치, 그리고 아름다움을 발견하게 될 것이다.

▶ 참고문헌

1. Margulies, M., Egholm, M., Altman, W. E., Attiya, S., Bader, J. S., Bembem, L. a, Berka, J., et al. (2005). Genome sequencing in microfabricated high-density picolitre reactors. *Nature*, 437(7057), 376-80.
2. Salzberg, S. L., Phillippy, a. M., Zimin, a. V., Puiu, D., Magoc, T., Koren, S., Treangen, T., et al. (2011). GAGE: A critical evaluation of genome assemblies and assembly algorithms. *Genome Research*.
3. Lin, Y., Li, J., Shen, H., Zhang, L., Papisian, C. J., & Deng, H.-W. (2011). Comparative Studies of de novo Assembly Tools for Next-generation Sequencing Technologies. *Bioinformatics (Oxford, England)*, 27(15), 2031-2037.
4. Nanopore genome sequencer makes its debut. <http://www.nature.com/news/nanopore-genome-sequencer-makes-its-debut-1.10051>
5. 클레이 셔키 저|이충호 역. 많아지면 달라진다. 갤리온 2011.