

# 생물학적 데이터 분석의 총아

글\_ 조승주 박사 KIST 생화학물질 연구센터 chosj@kist.re.kr

## 기획연재순서

- ④ 21세기의 수학
- ⑤ 21세기의 천문학
- ⑥ 21세기의 해양학
- ⑦ 21세기의 지질학
- ⑧ 21세기의 생태학
- ⑨ 21세기의 기상학
- ⑩ 21세기의 생명과학
- ⑪ 21세기의 생물정보학
- ⑫ 21세기의 우주학
- ⑬ 21세기의 고고학
- ⑭ 21세기의 인류학
- ⑮ 21세기의 생물분류학

**물**리학과 화학이 과학발전을 주도하던 시대는 가고 축적된 과학과 기술을 바탕으로 복잡하고 모호한 생명현상마저 과학적으로 접근할 수 있는 시대가 되었다. 십여년 전만 해도 생물학은 너무 복잡하여 근본적인 원인을 이해하기 어려웠지만 이제는 많은 부분을 보다 체계적으로 이해할 수 있는 기반이 마련되었다. 이러한 것들이 가능하게 된 근본적인 이유는 과학과 기술이 충분히 발전하였기 때문이다. 현재는 생물학적인 자료를 다량으로 축적하는 기술, 이 자료를 분석하여 정보를 이끌어 낼 수 있는 기술들이 개발되어 있다.

초기에는 인간유전체 사업 등에서 염기 및 아미노산의 서열과 같은 생물학적인

정보를 대량으로 저장하는 데이터베이스를 구축하고 유지하는 것이 생물정보학의 출발이었다. 그 이후에 이 축적된 자료를 보다 체계적으로 해석할 수 있는 많은 알고리즘의 개발이 병행되어, 현재는 염기 및 아미노산 서열, 단백질 도메인, 단백질 구조, SNP 연구, DNA-chip 해석 등과 같은 데이터를 분석하고 해석하는 것이 주요한 분야가 되었다. 기존의 학문과 다른 가장 큰 특징은 이 학문이 인터넷을 중심으로 발전하고 있다는 것이다. NCBI(미국국립생물정보센터) 등에서 인터넷으로 제공하고 있는 BLAST와 같은 서비스들은 실험생물학자들이 아무렇지도 않게 거의 매일 쓰고 있지만 사실은 방대한 자료와 분석 알고리즘의 총화이며 인터넷으로 접근이 쉽도록 개발된 중요한 생물정보학적 산물이다.

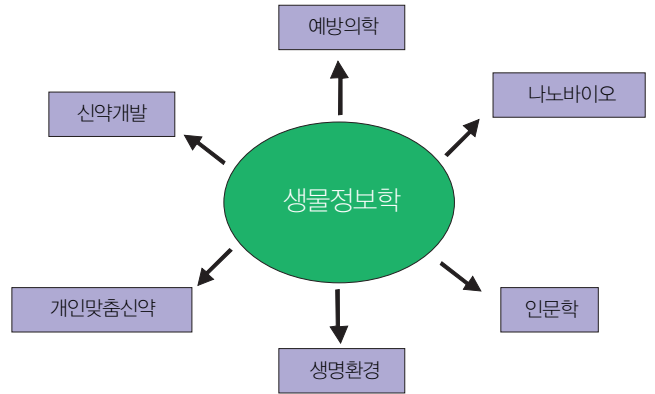
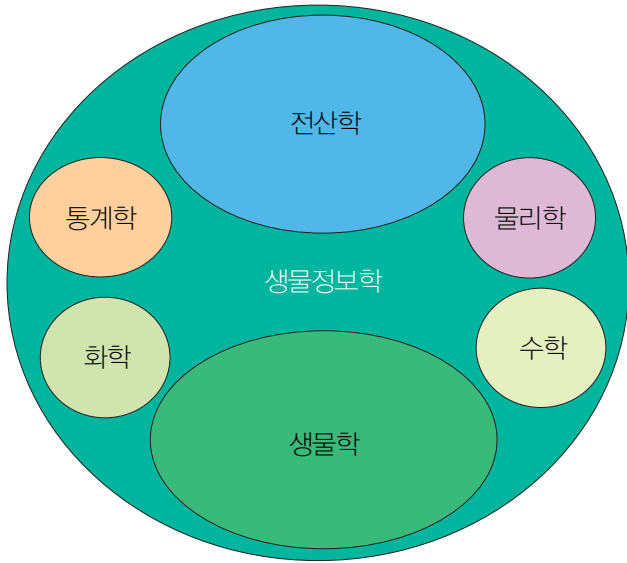
생물정보학이란 생물학적인 자료의 다양한 기술을 통한 축적을 돕고, 그 결과로서 얻어지는 자료에서부터 정보를 얻는 학문으로 정의할 수 있다. 이를 위해서는 생물학과 전산학 분야뿐만 아니라, 통계학, 화학, 물리학, 수학 등의 다른 학문분야의 지식도 필요한 다학제간의 영역이 융합되어야 하는 분야다. 현재 생물정보학의 발달과 함께, 분자생물학을 연구하는 실험실은 점점 더 대용량 데이터베이스와 다양한 분석도구 등을 연결할 수 있는 컴퓨터에 의존도가 높아지고 있다.

## 생물, 전산, 화학, 통계학 등 융합된 학문

생물정보학이 이렇게 응용범위가 넓고 그 가능성이 무궁무진하다는 것에 이의를 제기하는 사람은 많지 않을 것이다. 특히, 인간유전체사업을 통한 인간 염기서열의 해독이 성공한 이래 이를 활용할 수 있는 강력한 도구로서 생명정보학에 대한 관심이 촉발되었다. 한때 전세계적 생명정보 벤처회사들이 많이 생겨났지만, 성공한 회사는 지금 많지 않다. 이 중 일부는 큰 회사들과 전략적인 제휴를 통하여 사업모델을 세우고 있다.

대단히 많은 자료가 축적되었고 이를 활용할 수 있는 많은 알고리즘이 개발되었으며, 컴퓨터의 기능도 날이 향상되고 있고, 점점 더 많은 응용분야가 태동하고 있음에도 불구하고, 생명정보학은 아직도 걸음마 단계에 불과하다. 이는 생명현상이 애초에 생각했던 것보다 더 복잡해서 여기서 유용한 정보를 추출해 내는 것이 쉽지 않기 때문이었다. 다시 말하면 생물현상을 이해하기에는 여전히 우리가 얻는 지식이 충분하지 못하다는 것이다.

생물정보학을 통하여 신약을 개발하고 나아가 다른 인접 학문들과 융합하여 응용분야를 넓히는 것은 중요한 일이다. 하지만 그보다도 더 눈여겨 보아야 할 것은 이 학문이 가지는 인프라적인 가치이다. 이것은 마치 국토개발을 위해서 고속도로를 만드는 일과 같다. 고속도로를 만드는



생물정보학은 생물학과 전산학을 주축으로 물리학, 화학, 통계학, 수학 등이 필요한 융합학문이며, 여러 응용분야에 강력한 도구이다.

자체만으로는 직접적인 부가가치를 내기는 힘들 것이다. 다시 말해서, 고속도로 이용료를 징수하는 것만으로는 자체적으로 수지를 맞추기 어려울 수도 있다. 하지만 고속도로 개발을 통한 교통의 발달은 국가의 경제발전에도 크게 이바지하였으며, 이 부가가치가 말로 고속도로를 건설해야 했던 이유이다.

현재로는 생명과학이 어떤 분야에 크게 이용될 것인지에 대하여 말하기는 어렵다. 하지만 어떤 면으로든 반드시 중요하게 이용될 것이다. 그러면 이렇게 중요한 생명정보학을 어떻게 발전시킬 것인가? 이에 관련하여 생물정보학은 다학제간의 융합학문이라는 점이 중요하다.

실제로 전문가가 되기 위해서는 생물, 전산, 화학, 통계학 등의 많은 부분에 대한 상당한 지식을 요구한다. 이는 개인적인 측면에서는 장시간 끈기를 가지고 꾸준히 연마해야 되는 학문임을 의미한다. 나아가서는 국가적으로 전략적인 차원에서 육성해야 한다. 단기간에 가시적인 성

과를 요구하는 토양에서는 생물정보학이 자리잡기 힘들 것이다. 긴 안목을 가지고 전문가 교육문제부터 차근차근 구축해 나가야 하는 기초학문인 것이다.

**인문학에까지 응용되는 등 가능성 무궁무진**

현대에는 학제간의 벽이 허물어지고 있다. 기존에는 전혀 관계가 없는 것처럼 생각되었던 전산학과 생물학의 결합으로 생겨난 생물정보학이 그 대표적인 예가 될 것이다. 애초에 생물정보학은 신속하게 신약을 발견하고 검증하는 기능을 가질 것으로 생각되었다. 질병을 예방하는 분야와 개개인에 대한 맞춤 신약은 아직 활성화되지는 않았지만 미래에는 분명히 일부 가능해질 것으로 보인다. 그럼에도 불구하고 처음에 기대되었던 만큼의 속도 향상은 아마 당분간 힘들 것이다.

학제간의 통합은 미래에도 더 가속화 될 것이다. 나노바이오, 생명환경 같이 새롭게 생겨나는 융합분야에도 생물정보학은 여전히 중요한 역할을 할 것으로 기대

된다. 고고학에 물리화학적 기술인 방사선 동위원소 연대측정법이 도입되어 정확한 연대를 예측하는데 크게 공헌하고 있다. 이와 마찬가지로 생물정보학도 인간의 근원을 밝히는 인문학에까지도 그 응용범위를 넓혀가고 있다. 즉, 사람과 침팬지, 고릴라 등과 같은 생물의 유전체를 비교함으로써 ‘사람과 원숭이와 차이점’과 같은 인문학적 의문까지도 자연과학으로 분석하려 하고 있다.

또한 생물정보학에 중요하게 활용되는 많은 알고리즘들이 로봇을 만드는데 쓰일 수 있다. 생물정보학의 응용범위는 이처럼 무궁무진하다. 특히, 실험방법이 자동화되고 대량의 자료가 쉽게 축적됨에 따라 대량의 자료를 처리할 수 있는 생물정보학의 중요성은 더욱 증가할 것이다. ⑤



글쓴이는 성균관대를 졸업 후 포항공대에서 박사 학위를 받았다. UCSF 연구원, EPA 연구원, LG 생명과학 연구원을 지냈다.