Bioinformatics and Biosystems 2006, Vol. 1, No. 1, pp. 1-7

생물정보학과 환원주의 - 생물정보학은 Computational Molecular Biology와 동의어인가?

엄 융 의

서울대학교 의과대학 생리학교실

초 록

현대 생물학은 온통 정보의 홍수에 넘쳐나고 있다. 이렇게 쏟아져 나오는 정보들을 체계적으로 정리하고 이해하고 파악하는 것은 매우 중요하다고 하겠다. 생물정보학은 이렇게 쏟아져 나오는 정보들을 수학, 전산학, 정보학 등의 방법론을 사용하여 체계화시키려는 새로운 학문이고 미래 지향적 융합 분야이다.

Bioinformatics의 정의

바이러스로부터 미생물, 식물, 동물 그리고 인간에 이르기까지 수많은 유전체사업이 이미 완료되었거나 계획 추진되고 있다. 이러한 유전체 사업의 결과로 각종 생물체의 유전체 염기서열에 대한 정보가 쏟아져 나오고 있다. 이와 같이 컴퓨터를 이용하여 생명체 관련 자료를 체계적으로 정리하고, 분석, 이용하는 방법을 연구하는 새로운 분야가 탄생하게 되었고 이 분야를 생물정보학 (Bioinformatics) 또는 생명정보학이라 부르게 되었다. Bioinformatics는 90년대 후반부에 들어오면서 HGP (Human Genome Project)의 활성화와 DNA chip 등 제반 기술들의 발달로 인해 더욱 더 풍부한 Data들을 활용할 수 있게 되었다.

그렇다면 생물정보학이란 정확히 무엇을 하는 학문 분야일까? 또 그와 연관된 혹은 유사한 용어들은 어떤 의미를 내포하고 있을 것인가? 우선 필자가 이 글을 쓰는데 이 문제를 한번 짚고 넘어 가는게 중요하다고 생각되어 정리를 해 보았다.

넓은 의미로 생물정보학이란 컴퓨터를 사용하여 생물 정보를 다루는 여러 형태의 접근 방식이라고 정의할 수 있겠다. 그러나 좀 더 분명하게 정의하자면 컴퓨터를 사용하여 생체 물질의 구조와 조성에 대하여 정보를 저장, 비교, 추출, 분석하고 예측하는 분야라고 정의할 수 있다. 여기에 모델링하는 일 까지를 추가할 수도 있을 것 이다. 여기서 생체 물질이라 함은 주로 유전물질, 단백질을 말하며 고전적인 생물정보학은 여기까지를 말한다. 그래서 좁게, 분명하게는 " DNA, 아미노산 순열 및 그와 연관된 정보를 이용, 생명현상을 해결하는 것을 목적으로 하는 수학적, 통계학적, 컴퓨터 방법을 사용하는 학문분야" 라고 하기도 한다. 최근 들어 여기에 비교유전체학 (comparative genomics), DNA microarrays 기술, functional genomics, proteomics, structural genomics, medical informatics 등을 포함시키기도 한다.

유사 용어, 분야

그럼 생물물리학 (Biophysics)이란 분야와는 어떤 관계일가? 정의에 따르면 생물물리학이란 생물학적 구조와 기능을 밝히기 위하여 물리학적 방법을 사용하는 다학제 분야이다. 그렇다면 생물정보학과 상당부분 중첩되어 있다고 할 수 있다. 또 계산생물학 (computational biology)과 비슷한 의미로 쓰이는데 계산생물학이란 생물학 연구에 컴퓨터를 사용하여 접근하는 방식을 말하며 엄격하게 말하면 학문의 분야로 보기는 어려울지 모르겠다. 의료정보학 (Medical Informatics)은 생물정보학의 한 분야인가? 의료정보학은 의료정보 및 data 처리를 위하여 알고리즘과 구조에 관하여 연구하는 분야로서 의료정보 자체보다는 그것의 처리 방식에 더 중점을 두는 학문 분야이다. 그리고 현재의 생물정보학과 달리 세포 이상의 수준 (인구 집단에 이르기 까지)을 주로 다룬다. 화학정보학 (Cheminformatics)은 주로 신약 개발을 위하여 화학물질의 합성, 생물학적 screening, 자료 저장 및 추출에 관한 학문 분야라고 이해된다. 수리생물학 (Mathematical Biology)은 생명현상을 수학적 원리로 해석하려는 학문분야로서 문제를 해결하거나 컴퓨터 언어로 꼭 처리할 필요는 없다. 약물유전체학 (Pharmacogenomics)은 약물의 target을 확인하기 위하여 유전체학의 접근방식과 기술을 사용하는 학문분야이다. 따라서 약리학과 유전학과 관련된 정보를 정리하고 처리하기 위하여 생물정보학의 도구를 이용하는 분야라고 할 수 있다. 약물유전학 (Pharmacogenetics)은 약물유전체학의 한 분야로서 SNP, 환자의 약물반응 등 과 유전사이의 관련성 파악을 위하여 유전체/생물정보학적 방법을 사용하는 학문분야를 말한다. 이상과 같이 생물정보학이 내포하고 있거나 관련이 있거나 유사한 분야들을 간단히 살펴보았다. 어떻게 보면 차이가 확실하게 정립되는 것 같기도 하고 또 다르게 생각하면 여전히 혼돈 상태에서 벗어나지 못하고 있음을 발견하게 된다.

대부분의 정의에서 공통적인 점은 생물학적인 문제를

tackle하는데 있어서는 동일하나 아직은 생물학적이라기보다는 전산학, 수학, 물리학적인 테두리를 크게 벗어나지 않았다는 점이다. 그리고 20세기 후반 생물학의 주된 흐름인 환원주의적 사고에 머물러 있음을 알 수 있다. 아마 그것은 수학, 물리학, 화학, 공학 등의 기본적인 배경이 그렇기 때문에 피할 수 없을지도 모른다. 우리는 여기서 중요한 문제를 하나 짚고 넘어가야 한다고 생각한다. 생물정보학에서는 주 관심사가 생물학자들이 생산해내는 자료들을 어떤 도구를 써서 database화하고 어떤 방법으로 retrieve하는가에 있다고 할 수 있을지 모르나 필자 같은 생리학자들의 입장에서 보면 지금 언급한 사항들도 중요하지만 (특히 유전체, 단백체의 경우 특히 그러하지만) 어떤 data가 실지 생명체에서 일어나는 것인지 취사선택하는 것이 훨씬 중요하다고 생각한다. 왜냐하면 생체의 dynamic data는 기존의 발표된 자료들이 서로 상반되고, 특수 상황 하에서 얻어진 것이 많으니까….

이제 현대 생물학, 생물정보학의 환원주의적 접근방식에 대한 반성과 해결 방안에 대하여 생각하여 보기로 하자.

현대생물학에서의 환원주의

생명과학은 20세기 후반부터 현재까지 소위 환원주의(還元主義, reductionism)라는 극단적인 미시적 과학사상이 지배하고 있다. 특히 인간게놈프로젝트가 완성된 후 모든 생명현상, 질병의 원인, 치료 방법을 유전자에서 찾고자 하는 노력이 현대생물학의 대세를 이루고 있다. 그러나 인간게놈프로젝트를 완성한 과학자들은 결론에서 "인간이 찾던 유전자의 긴 사슬이 밝혀졌다. 그러나 이들 유전자 정보를 이해하는 것은 턱없이 부족하다" 라고 지적하였다. 이 말은 유전자의 사슬은 상당히 밝혀졌지만 그들이 의미하는 생리기능은 너무 모른다 라는 의미이다.

그런 지적에도 불구하고 많은 생물학자들은 유전자만능의 사고에 젖어있다. 특히 우리나라에서는 더욱더 그러하다. 그럼 왜 사람들은 유전자 만능주의에 빠져있고 일반인들은 특히 더욱 그러하다. 그건 아마 인간 두뇌의 특성 때문이리라. 인간은 두뇌를 통하여 사물을 관찰하고 이해한다. 인간의 두뇌는 대단히 막강한 사고능력을 가졌지만 지극히 선형적 (線形的) 사고에 치우쳐 있다. 특히 현대인은 모든 것을 단순화하는 선형적 사고에 잘 길들여져 있다. 비선형적 현상에 대하여는 우선 받아들이기 어렵다. 그렇다면 생명현상은 어떠한가?

유전자는 단백질을 만드는 암호코드를 가지고 있다. 그런데 유전자에 의한 단백질 합성은 사람들이 생각하는 것처럼 선형적 현상이 아니다. 인간이 가지고 있는 유전자수는 30,000개 정도라고 알려져있다. 그렇다면 유전자가 만드는 단백질은 몇 개나 될까? 선형적으로 생각하면 (많은 학자들도 그렇게 생각하여 왔다) 30,000개 이하래야 한다. 유전자 속에는 단백질을 암호 코드하는 신호 이외에 다른 정보도 가지고 있으므로. 그러나 세포 속에 존재하는 단백질은 20 만개에서 백만 개 정도가 되리라 생각되고 있다 (정확하게 숫자조차 파악 못하고 있는 실정이다). 이 사실은 단백질을 만드는 과정부터 선형적으로 설명을 할 수 없다는 것을 의미한다. 어디 이뿐이랴! 단백질들이 모여서 만드는 세포의 기능은 더더욱 비선형적이다. 세포가 모여서 조직을 만들고 조직이 모여 장기, 장기들이 모여 시스템, 이들이 결국 하나의 인간을 만드는 것이다. 이러한 지극히 비선형적 현상의 극치인 인간을 가장 단순하게 유전자와 유전자가 만드는 단백질의 기능으로만 해석하려는 현대생물학은 한계점을 보일 수밖에 없다. 따라서 우리는 이들의 해석과 실험결과를 경이의 눈으로 바라보고 정량화하고 database화하기에 급급하기에 앞서 그것을 검토하고 분석하고 판단할 수 있어야 할 것이다. 과연 생물학자들이 생산해 내는 자료가 모두 쓸 만한 자료인가를.

센트럴 도그마에서 벗어나는 사실들

에너지대사를 주관하는 미토콘드리아가 어머니를 통해서만 유전된다는 사실과 일맥상통한다. 이 사실은 상상을 하기 어려울 만큼 큰 의미를 가지고 있다. 지금 우리가 진실로 알고 있던 다윈 (Darwin)이래의 신 다윈 학설이나 20세기 최대의 발견으로 얘기되는 왓슨-크릭의 central dogma가 항상 성립되지 않음을 시사한다. 후천적으로 얻은 형질이 후대에 유전된다는 라마르크의 신 라마르크 학설이 다시 등장하게 될 상황에 이르렀다. 누구도 자신 있게 답할 수 없는 문제이지만 인류가 끊임없이 논쟁을 계속해온 유전자 (Nature, 씨)가 중요하냐, 환경 (Nurture, 교육)이 중요하냐의 명제에 되돌아온 것 이다.

유전자가 생명의 모든 것 인가?

환원주의자들은 인간 생명의 시작부터 그 진행 프로그램까지 모든 것이 유전자에서 온다고 주장한다. 20세기에는 대다수의 학자, 일반인들에게 그들의 주장을 설득하고 세뇌하는데 성공하였다. 그러나 아무도 아직 그 프로그램을 찾아내는 데는 실패하였다. 그렇지만 그들은 유전자 숫자가 단백질 숫자보다 적어 그것을 설명하는 것도 어려운 일로 등장하는 상황에서 지금은 이런 논쟁을 할 단계가 아니며, 환원주의적 접근만큼 합리적이고 과학적인 것이 없다고 말한다. 일견 그들의 주장은 매우 타당한 것처럼 보인다.

그렇다면 유전자 속에 생명 프로그램이 어떻게, 어디에 존재할 것인가? 시기와 경우에 맞게 생체기능을 변화시키려면 유전자에서 단백질을 만들어 보내야 가능하다. 그렇다면 유전자는 때에 맞춰 단백질을 다르게 합성할 수 있어야 한다. 유전자에서의 전사 명령은 거의 단백질이 주로 맡아서 한다.

그렇다면 그 단백질은 어디서 온 것인가 라는 의문이 제기될 수밖에 없다. 즉 닭과 계란 누가 먼저 인가 라는 원초적 모순에 접하게 된다.

우리 생활 속의 예를 좀 들어보자.

요즈음 웰빙이 우리에게 중요한 키워드가 되고 있고 웰빙의 일환으로 심장병을 예방한다고 알려진 적포도주에 대한 관심도 매우 높다. 캬베르네 소비뇽 (cabernet sauvignon) 이란 적포도주 품종을 한번 살펴보자. 이 포도를 원료로 하여 프랑스 보르도, 미국 캘리포니아, 호주, 칠레, 남아프리카 공화국 등에서 좋은 적포도주가 생산되고 있다. 잘 알려진 것처럼 캬베르네 적포주는 가격은 물론이고 맛에서도 천차만별이다 (유전자는 동일한데도). 이런 맛의 차이는 어디서 오는가? 미국 Napa valley에 있는 포도주 양조업자에게 물어보면 아마도 이렇게 말할 것이다. 포도주의 맛은 포도 종류, 기후, 영농법, 양조기술, 그리고 토양 때문에 달라 진다 (미국은 기술적 측면을 강조) 라고. 같은 질문을 프랑스 포도주 제조업자에게 해보면 그들은 한마디로 테로아 (terroir)의 차이라고 말한다. 테로아는 프랑스말로 흙의 의미를 갖는 단어로서 토양, 지역 및 국소 기후, 포도밭의 위치, 포도밭의 경사도 등 사람에 의한 기술적인 것을 제외한 모든 환경적인 것을 뜻한다.

사람은 원래 한 조상 어머니에서 유래했을 거라고 생각하고 있지만 지금 보면 그 말이 믿겨지지 않을 정도로 차이가 크게 느껴진다. 유럽 사람들을 예로 들어 보자. 북구 사람들은 햇빛을 잘 못 받아 피부가 희고 (햇빛 많이 받으면 피부병 잘 생김), 키가 크고(나무를 촘촘하게 심으면 모두 더 많은 햇빛을 받기 위해 키가 쏙쏙 자라듯이), 감정적으로는 침착, 냉정하고 어떻게 보면 좀 우울한 편이다(날씨 춥고, 긴 겨울의 영향일 듯). 스웨덴 출신의 소렌스탐 선수가 플레이하는 것을 보면 너무 분명해 진다. 한편 지중해 연안 사람(프랑스, 이태리, 스페인 사람들)은 햇빛 많이 받아 피부 약간 까무잡잡하고, 키가 좀 작고, 말이 많고, 감정적이고 다혈질이고, 놀기 좋아하고, 낙천적이다. 물론 이들은 북구와는 다르게 미식가들이다 (물론 모든 사람이 다 그렇다는 말은 아니다).

다른 예를 들어 보자. 요즈음 여학생의 초경은 초등학교 때에 일어난다. 예전에는 중학교 때에 일어났었고, 일제시대 하에 조사된 것을 보면 더욱 늦게 시작되었다. 이 문제를 유전자 매니아는 어떻게 설명할 것 인지 궁금하다. 유전자 속에 초경이 일어나는 프로그램이 들어있다면 말이다. 그것보다는 인간 생명체에 초경이 일어날 기반은 되어있는데 환경의 조건에 따라 (영양, 노동, 성에 관한 정보 등), 성적 자극을 덜 받으면 늦게 일어나고 일찍부터 많이 노출되면 빨라질 수밖에 없지 않을까? 다시 말하여 환경의 영향 하에 유전자의 활동이 조절됨을 시사하는 것이고 유전자의 활동이 세포, 장기, 개체, 즉 고위수준의 활동에 따라 조절됨을 말하여 준다.

또 다른 예를 들어보자. 얼마 전까지만 해도 우리나라 사람들에게 암은 위암, 간암, 자궁경부암이 주종을 이루어 왔다. 그러나 지금은 대장암, 폐암, 유방암이 더 많다. 유전자보다는 환경 변화가 중요하다는 증거이다. 이는 즉 환원주의의 원칙에 위반되는 것이다.

생명의 프로그램은 유전자에 쓰여 있는 대로 진행되는 것이 아니라 유전자가 세포, 장기, 개체 등 고위 수준의 활동에 따라 환경의 조절을 받으며 진행된다고 말할 수 있다.

환원주의자들 주장의 허와 실

우리나라 생물학자들 중 많은 사람들은 한국인에게 많은 질병의 원인을 유전자에서 찾아야 한다 라고 주장한다. 그리하여 한국인에게 많이 발생하는 암의 원인을 유전자에서 찾는 연구를 수행하여 여러 개의 유전자를 찾았다 라든지, 당뇨병 발병에 관여하는 유전자를 찾았다, 혹은 심장병을 일으키는 유전자를 찾았다, 노화에 관여하는 유전자를 찾았다, 치매를 일으키는 유전자를 찾고 이를 치료할 수 있는 획기적인 약물을 개발하였다, 암의 성장을 억제하는 획기적 물질을 발견하였다, 등의 톱뉴스가 매스컴을 심심치 않게 장식해 오고 있다. 일반 환자들은 이에 혹하여 행여 자기가 바로 혜택을 볼 수 있을 것 같아 가슴을 졸이기도 한다. 그런데 이게 사실이었으면 하는 바램을 필자 자신도 가지고 있지만 어디까지나 희망 사항일 뿐이다. 필자 개인적인 생각으로는 이런 주장이 나오고 보도까지 되는 데는 학자들의 책임이 크다. 복잡하기 그지없는 인체기능 조절 단백질 (적어도 20만개)중 1-2개의 단백질을 찾아냈다고 질병원인이 밝혀지고 생체 기능이 밝혀지고 치료제가 개발되었다고 말 할 수는 없다 (물론 아주 가능성이 없는 것은 아니지만). 그것을 가지고 모든 것이 해결된다고 주장하는 것 그 자체가 문제인 것이다. 우리 생체를 TV 수신기에 비유하여 얘기하면 좀 더 실감이 날 것 같다. 요즈음은 TV만드는 기술이 좋아 그런 경우가 드물지만 간혹 접촉 불량으로 화면이 나왔다, 안 나왔다 하는 경우를 경험해 보았을 것이다. 그러다 손으로 한대 치면 안 나오던 화면이 다시 나온다. 이 경우 아주 단순한 부품의 연결 불량이 화면을 안 나오게 한 것이다. 그렇다고 접촉 불량이 된 부품이 화면을 보이게 하는데 가장 중요한 부품이라고 (그럴 수도 있지만) 얘기할 수는 없을 것이다. 인간이나 TV (물론 인간에 비해 훨씬 간단하지만)는 복잡한 조절회로를 가지고 정교하게 기능을 하고 있기 때문에 부속품 1개, 혹은 유전자/단백질 한 분자가 어떤 기능을 맡아서 한다고 쉽게 결론 내릴 수는 없는 것 이다. 더구나 한 유전자가 한 단백질만을 만드는 것은 아니기 때문에 생명 현상은 우리가 상상할 수 없이 복잡한 것이다.

또 다른 예를 들어 보자.

필자도 20여년 동안 환원주의적 연구 방법으로 심장과 혈관의 생리학을 연구하여 왔다. 이 분야에서 중요한 문제 중의 하나가 심장의 자동능 (自動能, 외부 영향 없이 심장 스

스로 1분에 70번 가량 박동하는 능력)에 관한 것이다. 심장 세포에는 칼슘이온통로 단백질이 세포막에 존재한다. 이 통로를 약물로 억제하면 심장 박동이 멈춘다. 환원주의자는 사실을 근거로 칼슘통로 단백질이 심장 박동성을 결정하고 좌우하는 단백질이라고 주장한다. 심장 생리학자들은 이렇게 간단하게 말할 수 없다는 것을 잘 알고 있다. 근래 많이 보도된 새로운 발견이라는 것의 대부분이 이런 범주에 속한다고 볼 수 있다. 그러나 심장이 자동능을 나타내고 시의 적절하게 박동성을 조절하는 데는 적어도 100개 이상의 단백질이 관여한다고 알려져있다.

환원주의의 한계

20 세기 후반부터 꽃을 피우기 시작한 환원주의적 생명과학 연구는 인간 유전자의 구성을 밝히는 것을 비롯하여 그동안 해결하지 못하였던 여러 생물학적 문제를 밝히고, 질병의 진단과 치료에 획기적 진보를 가져오는 엄청난 공헌을 새웠다. 환원주의란 유전자는 단백질을 만들고, 단백질은 단독으로 혹은 단백질의 상호 작용에 의하여 모든 생체기능이 일어나고, 이런 단백질들이 모여 200 종류의 세포를 만들고 비슷한 세포 끼리 모여 조직을 만들고 장기를 만든다, 장기와 조직이 모여 시스템을 만들고 시스템이 모여 인간 개체를 만든다 라고 생각하는 현대 생물학의 거의 유일무이한 학설이다. 생명의 근원은 유전자에 있게 되고 이 유전자는 대대로 전파되어 비슷한 형질이 후세에 이어진다. 이런 과정 도중에 환경에 적응하는 과정을 거쳐 진화가 일어나 현재의 우리에 이르렀다고 보는 것이 바로 현대 생물학의 기본 사상인 셈이다. 이런 사고는 멘델, 다윈, 왓슨 · 크릭 등이 이어온 연구업적의 결과이다. 이 영향으로 20세기 후반의 생물학 연구는 유전자의 구조와 기능 (소위 유전체학, genomics)을 먼저 밝히고 그런 다음 단백질의 구조와 기능 (단백체학, proteomics)을 밝혀나가야 한다는 생각이 주류를 이루게 되었다. 단백질의 구조와 기능을 밝히고 난 후에 다음 단계로 넘어가야 된다는 생각인 것이다. 얼핏 지극히 당연하고 기초부터 차근차근 연구해 나가자는 주장에 누가 반론을 제기할 수 있겠는가?

생명체와 건물

환원주의자들은 인간 생명체를 건물에 비유하기를 좋아한다. 건물을 짓기 위하여서는 설계 (유전자) 대로 기초부터 차례 차례 진행되어야 한다고 말한다. 생명과학 연구가 건물 짓는 것과 같다면 당연히 그래야 한다. 그러나 생명체는 건물이 아니다. 건물 짓듯이 설계도가 있어 설계도대로 우리가 생명체를 창조해가는 것은 아니다. 생명체는 우리가 설계도대로 만든 것이 아니고 이미 존재해 우리 앞에 있는 것 이다. 마치 전혀 내용을 모르는 건물이 우리 앞에 있듯이. 미지의 건물을 사용하고 있는데 설계도가 없다. 재료가 무엇인지도 잘 모른다. 이럴 경우 우리는 건축 재료, 설계도를 알기 위하여 많은 노력을 해야 할 것이다 (유전자, 단백질에 대한 연구가 필요하듯이). 그러나 그것만으로는 충분하지 않다. 우리는 건물의 구조, 구성, 활용도 등 (해부학, 생리학 등)을 잘 알아야 하고 고장이 나면 근본적 원인을 알던, 잘 모르던 간에 수리 (임상의학, 치료의학)를 해서 정상적으로 활용해야만 한다. 그리고 그런 고장이 왜 생기고 그 원인이 무엇이고 어떤 방법으로 수리하는 것이 좋은지 (병리학, 약리학 등)도 연구해야 할 필요가 있다. 그리고 이들은 배타적이 아니고 상호 보완적이다. 이 점이 바로 환원주의의 미시적 시각만이 전부가 아니고 거시적, 통합적 시각에 의한 새로운 접근 방식이 필요한 이유이다. 바로 균형이 필요하다는 얘기이다.

유전자, 단백질 그리고 한자 (漢字)

인간의 모든 세포 (정자, 난자 제외)는 30,000개 정도의 유전자를 가지고 있고 이들은 20만개 이상의 단백질을 만들어야 한다. 1:1로 이루어진다 라고 믿어왔든 학자들에는 매우 실망스러운 사실이지만 유전자 하나가 여러 개의 단백질을 만드는데 관여해야만 설명이 가능하다. 그럼 어떻게 이런 일이 가능할 수 있을까? 유전자 1개가 여러 개의 단위로 구성되어 있고 이들이 경우 경우 마다 부분 부분 조합을 이루어 단백질을 만든다고 가정하면 가능하다 (실제 1개의 유전자는 여러 개의 단위로 구성되어 있다). 예를 들어 어떤 유전자가 A+I+B+I+C+I+D+I 로 구성되었다고 가정해보자. 전체가 다 참여하면 갑 이라는 단백질이 만들어 지는데 'A+I+D+I' 가 조합이 되고 다른 곳의 'Y+I'와 합해서는 '을'이라는 단백질을 만든다면 계산상 적은 수효의 유전자로도 더 많은 단백질을 만들어 낼 수 있다. 우리가 잘 아는 漢字를 가지고 유전자, 단백질 관계를 설명할 수 있다. 한자는 부수(部首)를 기본으로 하여 구성되어 있다. 보통 한자는 10,000자를 쓰는데 이들은 약 100개의 부수를 가지고 이루어져 있고 여기에 좌우, 상하, 내외, 본말, 주종을 이루는 성분이 추가되어 한자를 이루고 있다. 우리가 부수를 유전자로 한자의 單字를 단백질로 비유해 볼 수 있다. 한 부수는 여러 한자를 만드는데 사용된다. 예를 들어 보자. 告, 善, 喜 등은 '口' 부수를 공통적으로 가지고 있다. 만약 우리가 '善' 을 만드는 유전자가 있고 '善' 의 기능이 무엇인가를 알기 위하여 이 유전자를 파괴시켜 보는 실험을 해 본다면 (소위 유전자 knock-out 이라는 실험 방법). 우리는 '善' 의 기능을 알기 위하여 유전자를 파괴하였지만 '善' 이라는 글자에는 '口' 부수를 (다시 말하면 공통된 유전자 성분) 가지고 있기 때문에 '口' 부수를 기본으로 하여 만들어진 글자는 모두 영향을 받는 셈이다. 다시 말하여 유전자 하나는 어떤 한 단백질을 만드는 데만 사용되지 않고 여러 단백질을 만드는데 사용된다. 더욱 중요한 것은 인체의 여러 기능들 (심장 박동, 위 운동, 기억, 혈압 등…)은 하나의 유전자가 아닌 수십, 수백 개의 유전자 활동의 총합

으로 나타난다. 유전자와 단백질 사이의 관계는 많은 사람들이 생각하듯이 선형적 관계가 아니고, 단백질 - 세포 기능 사이, 세포 - 조직, 장기 사이, 장기, 시스템 - 인간 개체 사이의 관계는 더더구나 선형적 관계가 아니다.

새로운 접근 방법 - 피지옴(Physiome)

그럼 이런 비선형적 관계를 어떻게 접근하여 파악하고 이해하고 실제에 응용할 수 있을까? 경제학에 미시경제학과 거시경제학이라는 크게 다른 접근 방식이 있다. 경제 지표 하나의 변동 요인을 연구하고 이들이 전체 경제에 미치는 효과를 연구하는 분야를 미시 경제학이라고 하고, 개별 경제 지표보다는 여러 지표가 합쳐서 나타나는 전체 경제를 대상으로 여기에 영향을 주는 지표들을 찾아가는 것을 거시 경제학이라고 한다면 환원주의는 미시경제학과 유사하다. 미시 경제학으로 문제가 잘 안 풀리면 필연적으로 사람들은 거시 경제적 접근 방식을 취하여 문제를 해결하고자 한다. 생물학에서도 마찬 가지다. 미시적 환원주의 일변도의 접근방식에 한계가 드러난 이즈음 거시적 통합생물학의 등장은 필연적이다. 미시적 접근에서 벗어나 생명현상을 거시적, 통합적으로 보자는 움직임이 얼마 전부터 시작되었다. 그 전에도 거시적 시스템적 접근 방식은 존재하여 왔다. 그러나 미시적 지식이 거의 없었던 시절이었고 실험방법의 제한이 많아 20세기 후반에 들어서 환원주의 화려한 등장과 함께 그 움직임이 보잘것 없었다. 그러나 지금은 미시적 지식이 많이 축적되고 새로운 기술 (컴퓨터, CT, MRI, PET 등)이 발전하여 통합적 접근이 가능하게 되었다. 물론 환원주의적 접근의 한계성이 가장 중요한 촉매제이었긴 하지만.

이리하여 1998년도 세계생리학회에서 인간 피지옴 프로젝트 (Human Physiome Project)를 시작한 것을 필두로 전 세계적으로 거시적 통합생물학이 태동하게 되었다. 피지옴 (Physiome) 이란 용어는 1990년대 중반에 만들어진 말로 physio (생명을 뜻하는 라틴어, life)와 - ome (전체를 뜻하는 접미어로 genome의 ome과 같음)이 합쳐져 만들어진 말이다. 그럼 무엇을 뜻하는가? 피지옴이란 유전자, 단백질, 세포, 장기 등 여러 수준에서 얻은 지식들을 모으고 이를 통합하여 인체 기능을 거시적으로 파악하기 위한 새로운 정보의 데이터베이스를 말한다. 말은 간단하지만 엄청난 정보들을 모아서 통합적, 거시적으로 파악하고 이해한다는 것은 쉬운 일이 아니다. 피지옴 연구의 방법은 컴퓨터가 빠질 수 없다. 인간의 두뇌는 우수하지만 단점을 많이 가지고 있다. 우선 인간의 두뇌는 비선형적 지식을 통합하는데 강하지 못하고 정보 저장량에 있어서는 너무 한계가 있다. 이런 이유로 우리는 막강한 IT기술을 이용한다. 컴퓨터와 데이터베이스 기술을 이용하여 생체정보를 정량화하고 저장한다. 그리고 그런 자료와 수학, 물리학, 전자공학, 생물학, 기초의학, 임상의학 등을 이용하여 컴퓨터에 세포, 장기 모델 (가상 세포, 가상 장기)을 만든다. 이런 작업이 진보가 이루어지면 컴퓨터상에 가상 인간을 만들 수 있다. 가상 세포, 가상 장기, 가상 인간은 요즈음 논란이 되고 있는 생명윤리에 저촉될 걱정도 없다. 피지옴의 최종 목표는 유용한 가상 인간을 만드는 것이다.

이렇게 보면 피지옴은 생물정보학의 중요 접근 방식을 이용한다. 시스템 생물학, 가상 생물학 (in silico biology, e-biology) 이라고 부르기도 하며 계산 생물학, 생물정보학의 중요한 tool을 이용하고 있다. 그러나 많은 생물정보학자들은 simulation, modeling 분야를 생물정보학의 한 분야로 생각하지 않는 듯 하다. 필자의 견해로는 현재의 생물정보학이라는 bibliographic 한 구조에 호흡하며 박동하는 생명력을 불어 넣을 수 있는 분야가 바로 Physiome이라고 생각한다.

피지옴 연구의 내용 및 범위:

유전자, 단백질, 대사, 세포의 분화, 운동, 분화 등 세부 기능, 세포, 조직, 장기, 시스템, 인간 개체 모든 수준에서의 생체 현상이 피지옴 연구의 대상이 될 수 있다, 시스템 생물학자 들, 생물정보학자들은 여전히 환원주의적 원칙에 충실하여 유전자부터 시작하는 상향식 (bottom-up) 접근 방법을 선호한다. 그러나 앞에서 언급한 것처럼 상향식, 하향식 (top down) 원칙에 구애 받을 필요가 없다. 어느 수준에서든지 시작하면 된다. 세계생리학회에서 제시하는 인간 피지옴 프로젝트에서는 세포, 조직 장기 기능이 포함된 거시적 시스템을 단위로 하여 구성되어 있다. 다시 말하면 인간 생체에 관한 어떤 부분이건 어떤 수준이건 모두 피지옴적인 접근이 가능하다. 그러나 현재 만들어진 피지옴 모델은 정립된 database가 아니다. 모델에 따라 도입한 데이터가 다르고 식이 다르다. 유전자은행처럼 쌓아 놓으면 되는 정보들이 아닌 것이다. 그렇기 때문에 여러 종류, 여러 연구자에 따라 다른 모델이 만들어 지고 있다.

피지옴의 기대효과 및 응용

새로운 약이 개발되기까지는 평균 약 14년의 시간이 소요되고 연구 개발 비용 또한 17억불이 필요하다. 그럼에도 불구하고 신약으로 성공하는 것은 1% 에도 못 미친다. 신약 개발에 소요되는 시간과 경비를 절감하기 위한 필사적 노력이 계속되고 있다. 제약회사들에게 희망을 주는 연구가 바로 피지옴 연구이다. 최초의 물질 디자인부터, 흡수, 대사, 배설 연구의 약리학적 연구, 전임상, 임상 시험의 모든 단계에서 가상 세포, 가상 장기를 사용하여 시간, 비용의 절감을 가능하게 할 수 있게 된 것이다. 이 결과 실험동물을 사용하여 실험해야 하는 필요성이 감소되고, 인체시험을 할 범위가 크게 줄일 수 있다. 사회 윤리적 문제 또한 줄일 수 있다. 미국 식약품안전청 (FDA)도 피지옴 모델을 이용한 가상시험을 적극 권장하고 있다.

21 세기의 큰 도전 중의 하나는 로봇 기술의 응용이다. 의학에서는 로봇이 외과의사 대신 수술도 해 주고 위험한 작업을 대신 해 줄 것으로 예상하고 있다. 로봇이 정교하게 움직

이려면 사람의 움직임을 정확하게 입력 해 주어야 한다. 이렇게 생명 현상을 컴퓨터 프로그램화하는 작업이 피지옴 연구이다. 로봇의 발달을 위해서는 근육 운동 피지옴 연구가 필수적이다.

현재 우리나라 사람들이 희망을 갖고 자랑스럽게 생각하는 줄기세포의 연구도 마찬가지다. 지금은 난치병 치료가 전면에 등장하고 있지만 사실은 더 많은 연구가 필요하다. 간세포가 왜 다른 세포로 되지 않고 간세포가 되는지 우리는 아직 잘 모른다. 줄기세포의 특징은 여러 종류의 세포로 분화될 수 있다는 것이다. 우리에게 이로운 방향으로 분화될 것인지 아니면 해로운 방향 (예를 들어 암세포로)으로 분화될 것인지 아직은 우리의 능력으로는 조절이 어렵다. 따라서 중간 단계의 많은 연구가 필요하다. 심장박동이나 뇌기능 등의 고위 기능들은 다른 세포들과 동일한 유전자를 가졌지만 실제에서는 약 10,000 개 정도의 gene이 express 되어 있을 거라고 짐작하고 있고 많은 부분의 유전자가 여러 장기에서 반복 사용되는 것으로 추측하고 있다. 세포 종류 별로 S/W on 된 gene, S/W off 된 gene이 다르게 조합되고 일생 동안 시기별로 express되는 pattern 또한 다를 것으로 추측하고 있다. 그러나 우리는 그 비밀을 아직 잘 모르고 있다. 이런 상황 하에서 중간에서 다리 역할을 하는 연구가 필요하며 그 중에서 가장 전망 있는 연구는 피지옴 연구라고 생각한다.

생명 현상을 컴퓨터에 프로그램화하면 생명현상의 이해가 지금보다 훨씬 용이해진다. 따라서 정상 기능의 이상에 의하여 발생하는 질병의 원인을 해석하고 규명하는 일 또한 쉬워진다. 피지옴은 의료전달 체계를 혁신적으로 변화시킬 수 있다. 개개인의 체질상의 특성, 유전정보, 질병 기록, 검사 결과 등 모든 정보가 컴퓨터기록으로 저장되고 휴대용 저장 장치에 기록되어 세계 어느 곳에 가서도 자신의 의료정보를 의사에게 보일 수 있다. 치료 약물에 대한 정보도 저장하여 종합하면 꿈에 그리던 맞춤의학의 시대를 만들어 낼 수 있다.

세계 동향

2003년 9월 세계적 환원주의의 메카라고 할 수 있는 하바드 의대에서는 새로운 학과를 창설하였다. 학과 이름은 시스템 생물학과. 수천 만 불의 경비를 투자하고 수학, 물리학, 공학, 컴퓨터공학, 생물학자, 기초의학자, 임상의학자를 망라하는 구성을 가진 학과를 만들겠다고 발표하였고 다음과 같이 학과 창립의 필요성을 정리하였다.

"미시적 조각에 대한 연구를 많이 하면 할수록 생체 기능이 전체적으로 어떻게 이루어지고 있는지를 모른다는 사실이 더욱 분명해진다. 우리가 분자 생물학에서 얻은 지식들이 어떻게 세포를 구축하고 기능하는지, 또 세포들이 어떻게 상호작용하여 장기를 이루는지 이해할 필요성이 절실하기 때문에 문제 해결을 위하여 생리학적 예측 모델을 만드는 것을 최종 목표로 삼았다." "많은 신약 후보들이 실제 환자에서는 효과가 없다는 것에 대하여 우리는 이해를 잘 못하고 있었습니다. 이런 문제를 해결하기 위하여 예측 가능한 가상 세포, 가상 장기 모델을 만들고 나아가 가상 인간을 만들고자 합니다. 이를 위해서 새로운 학과를 만들기로 결정하였습니다".

한편 일본에선 문부과학성이 선도과제 (leading project)의 하나로 세포・생체 시뮬레이션 프로젝트를 2003년부터 시작하였는데 5년간 약 1600억 원을 투자하여 게이오대학, 교토대학, 고베대학이 응용분야 센타를 맡고, 기반 기술은 도쿄대학, 오사카대학, 나고야대학이 분담하여 연구를 하고 있다. 이 프로젝트의 총괄책임은 安西祐一郎 게이오대학 총장이 맡고 있다. 그는 이 프로젝트를 다음과 같이 요약하고 있다.

"우리들의 행복의 원천인 생명에 대한 지식을 창출해 내고, 인간의 생명에 관련된 모든 문제를 해결해 나가는 것은, 인류공통의 소망입니다. 본 프로젝트는, 생체 시뮬레이션의 성공 예를 구체적이며 명쾌하게 제시함으로써, 의학・공학・생물정보학의 융합분야에서 새로운 지평을 열어 선도적 역할을 하고, 인류의 행복에 대하여 크게 공헌해 나가는 것을 그 사명으로 하고 있습니다. 생명현상에 관련된 대사나 정보전달의 본래의 반응을 컴퓨터상으로 재현하고 세포・생체 시뮬레이션을 구축해낸다면, 생명과학의 기반연구에 있어서 실험가설의 입증을 효율적으로 이루어내는 것뿐만 아니라, 신약개발 등에서 필요한 스크리닝시험의 효율화・경비절감을 가져올 수 있습니다. 지금까지 일본은, 이러한 연구영역에서 세계를 앞서 나갔지만, 구미의 추격도 점점 거세지고 있습니다. 본 프로젝트는 [경제 활성화를 위한 연구개발 프로젝트]의 일환이면서, 산학연 연대에 의한 장래의 산업화 등을 구체적 목표로 설정하고 있습니다. 더욱이, 앞으로 대규모의 과제로 확대되어 투자해야 하는 융합과학 과제라는 점에서, 또 여러 다른 분야의 연구자가 융합하여 국제경쟁력을 가지는 연구체제의 확립을 지향하고 있습니다. 또한, 젊은 연구자와 중견 연구자에 의한 통합적 연구 수행을 장려하고, 일본이 계속적으로 이 영역을 리드할 수 있도록, 인재의 육성에도 중점을 두고 있습니다."

뉴질랜드와 영국은 e-science의 일환으로 암, 심장 등의 주요 부분 연구에 착수하였고 이를 위한 융합과학자 양성을 위한 대학교육 과정의 개혁을 추진하고 있다. 독일은 간 (肝)세포 연구로 집중하여 엄청난 연구비를 투자하여 연구를 하고 있다.

결 론

이제 생명과학은 중요한 기로에 서 있다. 지금까지 경주해 오던 환원주의적 미시 생물학 일변도를 타파하고 거시적 통합적 생물학을 과감히 도입하여 21세기의 생명과학을 이끌어 가야만 하는 시점에 와 있다. 생체 실험 위주의 연구에서

컴퓨터를 이용한 시뮬레이션 모델 구축과 모델 실험에 많은 역량을 쏟을 때가 되었다. 그리하여 좁은 시험관속 단백질이나 배양세포의 단백질 발현에 근거한 해석으로부터 좀 더 높은 수준 (세포 보다는 조직, 조직 보다는 장기, 장기 보다는 개체)에서 생명현상을 관찰하고 분석하고 해석 응용하는 통합 생물학, 통합적 생물정보학으로 변환을 해 나가야 할 것이다. 지금까지 실험실에서 생산해 내던 작은 생명과학 정보가 앞으로는 중요한 지적 재산으로 피지움 분야에 이용될 것이다. 하나 하나의 연구 결과는 그 가치를 알기 어렵지만 이들이 모여 통합되면 큰 상업적 가치를 창출해 낼 수 있다. 피지움 연구는 동물 실험, 인체시험을 최소화하기 때문에 윤리적으로도 매우 바람직한 방향의 연구이다.

작은 나라 뉴질랜드의 오클랜드 대학 의공학과는 50 여명 연구자들이 한 마음으로 가상 모델을 만드는 작업에 뛰어들어 세계 최고가 되었고 일찍이 영국 노팅햄 대학 물리학과는 전 연구원들이 핵자기공명장치 (NMR)을 인체에 적용하는 연구에 전력을 쏟아 오늘의 MRI 영상법을 발명해 내었다. 우리가 교훈으로 삼아야 할 모델이라고 생각된다.

이제는 생물정보학 분야도 환원주의 일변도의 접근 방식과 database library 식의 정적인 고정 틀에서 벗어나 역동적인 생명과학의 한 분야로 나아가 통합생명과학의 중요 선두주자로서 새로운 영역을 개척해 나가야 한다고 생각한다.