

Genomic Information을 이용한 신약 개발 연구

김 양 석

(주)이즈텍 대표이사

유전체학을 기반으로 하는 분자생물학의 급속한 발전은 생명체에 대한 인간의 이해정도를 급격히 향상시키는 동시에 생물 산업 전반에 걸쳐 많은 영향을 미치고 있다. 특히 신약 개발 연구는 가장 많은 영향을 받고 있는 산업으로써, 유전체 기반 신약 개발(genomics-based drug discovery)이라는 새로운 접근 방법이 점차 확산되고 있다. 또한 각 개인의 유전체의 차이 및 발현 양상의 차이를 밝혀 낼 수 있는 생물학 실험 기법들이 개발됨에 따라 먼 미래의 꿈으로만 여겨졌던 개인 맞춤형 의약(personalized drug)의 개발이 약물유전학(pharmacogenetics) 연구를 통해 현실화되고 있다.

유전체 기반 신약 개발

특정 질병에 의해 유도되는 생체내의 유전자나 단백질의 변화를 알지 못하는 경우 신약 개발은 정방향 접근(Forward chemical genetics), 즉 인간 병세포주나 질환 동물 모델을 먼저 만든 후, 질환을 극복할 수 있는 화합물질을 무작위 처리를 통해 찾아내는 방법이 주류를 이루었다. 이 경우 생체내의 목표로 하는 분자 타겟(molecular target)에 대한 정보가 없으므로, 화합물질의 개량이나 부작용에 대한 예측을 하기가 어렵다. 더욱이 무작위 처리를 통해 질병을 극복하는 화합물을 찾아야 하므로 시간과 경비측면에서 매우 비효율적인 접근 방법이라 볼 수 있다.

하지만 유전체 정보를 활용할 경우 역방향 접근-질병 메커니즘의 이해를 통해서 우선 분자 타겟을 정확히 선정해 내고, 그 타겟의 발현양상을 조절 할 수 있는 화합물을 찾아서 접근하는 방법-을 통한 신약 개발이 가능하므로 정방향 접근의 단점들을 보완할 수 있다[1]. 더욱이 최근 급속도로 발전하고 있는 화학 정보학(cheminformatics)의 가상 검색(virtual screening)등의 방법은 타겟의 발현을 조절할 수 있는 화합물들을 미리 컴퓨터 상에서 검색해 냄으로써 신약의 발굴에 소요되는 시간과 경비를 혁신적으로 줄일 수 있을 것으로 기대되고 있다.

2000년 인간 유전체 지도가 공개될 시점에 신약 개발에 활용

되고 있었던 분자 타겟의 개수는 500여 개 정도이다. 활용되고 있는 분자 타겟을 단백질 군별로 다시 분류하면 130여 군으로 나눌 수 있으며 그 중 G-protein coupled receptors(GPCRs), serine/threonine과 tyrosine protein kinase, zinc metalloproteinases, serine proteases, nuclear hormone receptor, phosphodiesterase의 6개의 군이 반 이상을 차지한다. 또한 아직 활용되고 있지는 않지만 알려진 분자 타겟과 공통된 활성 부위를 가지고 있는 후보 분자 타겟(druggable genome)은 3,051개 정도로 예측되며 그 분포는 그림 1과 같다. 유전체 정보를 이용한 분자 타겟의 발굴이 아직까지 초기인 점과 현재 유전체학의 발전 속도를 고려하면 향후 2~3년 내에 새로운 분자 타겟의 개수는 10,000개를 넘어 설 것으로 예측된다[2,3]. 이러한 분자 타겟의 대량 확보에는 DNA chip과 같이 생체내 유전자들의 발현 양상을 동시에 관측할 수 있는 실험 기술의 발달과 기하 급수적으로 증가하는 생물학 정보들 속에서 유용한 정보들을 창출해 내는 생물 정보학(bioinformatics) 기술의 발달이 원동력으로 작용하고 있다.

DNA chip의 경우 특정 조건에서 대량의 유전자들의 발현 정도를 정량적으로 측정 해 낼 수 있게 함으로써 실험 유전체학의 혁신을 가져왔다[4,5]. 더욱이 실험 방법의 지속적인 개량을 통해 실험 결과의 정확도는 높아지고 있는 반면 실험 비용은 점차 낮아지고 있어, 단기간 내에 유전체학 실험에 광범위하게 활용될 것으로 예측된다. 하지만 DNA chip은 생명 현상에 직접적으로 관여하는 단백질들의 발현 양상을 간접적으로 관측한다는 실험적 약점을 가지고 있다. 즉 유전자의 발현 양상과 단백질의 발현 양상이 항상 직접적인 연관성을 가지지 않는다는 점과 번역후 변형(post-translational modification)에 의해 활성이 조절되는 단백질의 경우 유전자의 발현 양상이 중요한 정보를 제공하지 못한다는 점등은 DNA chip을 통해 산출되는 실험적 결과의 신뢰성에 의문점을 제기한다. 따라서 최근에는 대량의 단백질들의 발현 양상을 직접 관찰할 수 있는 단백질 칩의 개발을 위한 많은 연구가 진행되고 있다[6,7].

생물정보학은 이미 유전체 연구를 위한 필수적인 기술으로써, 생물학 연구 전반에 걸쳐 활용되고 있다. 또한 초기의 단순한

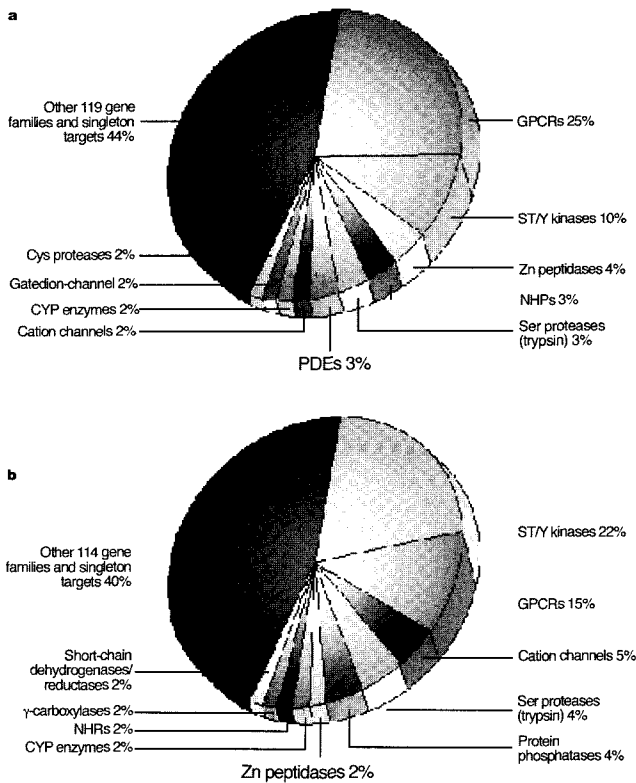


그림 1. 분자 타겟 단백질 군. a. 현재 활용되고 있는 분자 타겟 단백질들의 기능별 분류 b. 현재 활용되고 있는 단백질군의 특징을 바탕으로 분류한 후보 분자 타겟 단백질 군. 그림 출처:Hopkins, LA., Groom, CR. The druggable genome. *Nature Review Drug Discovery*. 1 727-730 (2002).

데이터의 저장이나 분석의 역할에서 벗어나 생물학자들이 실험을 통해 발견 할 수 없는 새로운 지식의 발견을 위한 데이터 마이닝에 중점을 두고 발전해 나가고 있다. 신약 개발을 위한 분자 타겟으로 많이 활용되고 있는 protein kinase의 경우 1,000개 이상이 생체 내에 존재하는 것으로 예측되어지고 있는데, 그 중 400여 개 정도의 단백질 발굴을 위해 지난 10여 년 간 전 세계의 많은 생물학자들이 지속적인 연구를 수행하여 왔다. 하지만 통계적 모델링 기법을 바탕으로 한 생물 정보학 기술은 이미 밝혀진 400여 개의 단백질들의 특징 분석을 통해 인간 유전체에 숨겨진 1,000여 개의 protein kinase 후보 단백질들을 불과 몇 개월만에 다 밝혀 내었다. 또한 최근에는 분자 타겟을 선별하는데서 그치지 않고, 타겟 단백질 내에서 활성에 영향을 주는 핵심 구조 및 서열을 예측하여 신약 개발을 위한 좀 더 정확한 정보를 추출해 내는 연구도 많이 진행되고 있어 향후 생물 정보학은 신약 개발에의 활용 및 기여 범위가 점점 더 확장될 것으로 예측된다[8].

하지만 이러한 역방향 접근법의 경우 병목이 되는 부분은 확보한 분자 타겟을 다시 실험 방법을 이용하여 재검증하여야 한다는 점이다. 즉 유전체 정보를 활용하여 대량의 분자 타겟

들을 빠른 시간 내에 검출 해 내는 것들은 가능하지만 각 타겟에 대해 실험적으로 검증하는 작업들이 필수적이며 현재까지는 이러한 검증 작업을 대량으로 정확하게 수행할 수 있는 생물학 기법들이 개발되지 않아 역방향 접근법의 병목으로 지적되고 있다.

이러한 병목 현상을 해결하고 분자 타겟의 정보를 신약 개발에 좀 더 직접적이고 효율적으로 응용하기 위해 최근에는 생물학과 화학이 결합된 연구 방법들이 많이 개발되고 있으며 대표적인 연구 분야를 화학 유전체학(chemogenomics)으로 생각 할 수 있다. 화학 유전체학은 유전체 정보, 신약 후보 물질 활용 가능성이 높은 저분자 화합물 라이브러리(small-molecule library), 그리고 화학 정보학의 기술들을 적절히 조합하여 효율적으로 신약 후보 물질을 찾아 내는 학문을 통칭적으로 의미한다[9]. 이러한 접근 방법의 경우 확보된 분자 타겟을 검증하는 작업이 완전히 없어지는 것은 아니지만 신약개발의 후반부 단계들을 좀 더 적합화 시킴으로 인해 전체 신약 개발 과정의 효율성과 성공률을 높일 수 있는 접근법이라 볼 수 있다.

역방향 접근법의 또 다른 큰 장점은 새로운 형태의 약 개발이다. 종래의 약들은 합성 화합물이나 천연 추출물이 주를 이루었던 반면, 유전체 정보를 활용하여 분자 타겟을 확정시킨 경우 생체 물질을 이용하여 분자 타겟의 활성을 조절할 수 있는 새로운 접근 방법들이 가능하다. 그 대표적 접근 방법으로는 anti-sense를 이용한 방법을 들 수 있고 이미 HIV, 파필로마 바이러스(Papillomavirus) 등 몇 개의 신약 후보 제품들이 임상 평가를 진행중인 것으로 보고되고 있고 최근에는 RNAi(RNA interface) 방법을 이용한 시도들도 많이 이루어지고 있다[10, 11, 12].

약물 유전학(Pharmacogenetics)

개개인의 약물에 대한 반응 차이를 분석하는 약물 유전학의 급속한 발전은 멀지 않은 미래에 개인 맞춤형 신약 개발의 시대가 도래할 것을 예측 가능케 한다[13]. 약물에 대한 개개인의 반응의 차이는 개개인의 유전체의 차이를 통해 많은 부분이 설명 될 수 있을 것으로 고려되고 있으며 따라서 유전체속에 포함되어 있는 SNP(Single Nucleotide Polymorphism)를 발굴하여 질병 혹은 약물과의 연관성을 분석하는 연구들이 활발히 진행되고 있다. 특히 대규모의 인력과 자금의 투자가 필요한 SNP 발굴 작업은 국제 사회에서의 공동 연구로 많이 진행되고 있으며 대표적인 컨서시움으로는 10개의 제약회사와 The Wellcome Trust가 함께 참여하여 300,000 개의 SNP 발굴 및 유전체 지도상에서의 위치 확인을 목적으로 하고 있는 TSC(The SNP Consortium)을 예로 들 수 있다[14]. 또한 GlaxoSmithKline을 비롯하여 Orchid Bioscience, Sequenom,

Illumina 등의 바이오텍들은 TSC에서 나온 결과를 기반으로 유전체를 검색할 수 있는 SNP marker를 제작하는 대규모의 공동 연구를 진행하고 있다.

현재까지 밝혀진 연구 결과를 바탕으로 임상 처방에서 개개인의 유전체의 차이 혹은 유전자의 발현 양상을 활용하는 시도는 이미 이루어지고 있으며 그 대표적인 경우는 trastuzumab (Herceptin)을 예로 들 수 있다. Trastuzumab은 인간화된 단일항체(humanized monoclonal antibody)로 HER2 암유전자의 발현을 조절하여 유방암의 치료제로 활용되고 있다. Trastuzumab의 처방은 HER2 암유전자의 발현 양상에 의존하는데, HER2 암유전자의 발현이 많이 되지 않은 2/3 정도의 유방암 환자에게는 처방이 의미가 없기 때문이다[15]. 다른 예로는 CCR5 cell-surface chemokine receptor를 분자 타겟으로 하는 AIDS 약그룹을 들 수 있다[16]. 이 그룹에 속하는 약들은 432 변종을 가지고 있는 AIDS 환자군에는 약의 효과가 없는 것으로 보고되고 있고, 이것은 바이러스의 유전체 타입과 약물의 직접적인 연관성을 보여주는 결과로 볼 수 있다.

현재 발굴되고 있는 SNP들은 계속해서 약물의 처방에 활용될 것으로 예측되고 있으나 현재까지도 유전체의 차이와 약물의 효과와의 직접적인 연관성을 밝히는 일은 쉽지 않다. 그 이유는 새로운 SNP를 찾기 위해 소요되는 시간과 경비의 문제와 함께 질병 메카니즘의 복잡성을 들 수 있다. 즉 대부분의 경우 질병의 발생 혹은 치유에 관련된 생체 과정은 한 개의 단일 유전자에 기인하기보다는 여러 개의 유전자들이 동시에 관여하는 경우가 일반적이기 때문이다. 따라서 유전체 변이와 약물과의 연관성도 단일 유전자 혹은 단일 변종으로 해석되기보다는 다중 유전자 변이와 약물간의 상관 관계를 밝히는 것이 필요하다. 따라서 생체내의 네트워크에 대한 이해가 약물유전학과 병행되어야 하고 관련 분야를 연구하는 메타볼롬(metabolome), 피지옴(physiome) 등의 분야가 유전체 연구의 새로운 중요한 분야로 인식되는 것도 같은 맥락에서 이해될 수 있다.

약물 유전체학(pharmacogenomics)과 독성 유전체학(toxicogenomics)

유전체학은 신약 개발의 직접적 활용 및 처방이외에도 약물의 생체 내 전반적인 활동 기작에 관련된 정보를 얻기 위해 활용되고 있으며 대표적인 연구 방법으로 약물 유전체학과 독성 유전체학을 들 수 있다. 약물 유전체학은 약물 유전학과 용어상에 혼동을 줄 수 있으나 그 의미는 구분되어 사용된다. 위에서 언급한 바와 같이 약물 유전학은 개개인의 유전적 변이를 기반으로 한 개개인의 약물의 반응에 중점을 둔 반면 약물 유전체학은 유전체 수준에서 약물이 유전자의 발현 양상이나 단백질들의 발현에 미치는 영향에 중점을 둔 학문이다. 독성 유

생물산업

전체학도 유전체의 발현 양상에 중점을 둔다는 점에서 약물 유전체학과 유사한 개념으로 이해될 수 있으나 약물의 생체 내 활성보다는 생체 내 독성에 중점을 둔다는 것이 차이점으로 지적될 수 있다. 유전체를 구성하는 전체 혹은 대부분의 유전자들의 발현 양상을 동시에 관찰할 수 있는 DNA chip 기술은 약물 유전체학과 독성 유전체학의 핵심 기술로 많이 활용되고 있으며 약물 유전체학의 경우 이미 많은 질병 및 관련 약물 연구에 활용이 되고 있다[17,18,19]

독성학은 신약 후보 물질의 개발 여부의 결정 및 처방 방법의 결정 등 신약 개발 과정의 플랫폼 기술로서 중요한 의미를 가진다. 현재 임상 혹은 전임상에서의 독성 판정은 주로 세포 해부학적인 방법들이 많이 적용되고 있으나 측정 가능한 한정된 부분만의 관측이 가능할 뿐 생체내의 전반적인 유전자 혹은 단백질들의 발현 양상을 측정할 수는 없다. 따라서 독성 판정의 실패 확률은 신약 개발의 기간 단축과 비용 절감에 위협 요소로 작용하고 있다. 최근의 유전체 실험 기법을 이용한 독성 판정 시도는 이러한 기존 방법의 보완 혹은 대체 가능성을 제시하고 있다. 특정 약물의 독성을 DNA chip을 이용한 전반적인 유전자 발현 양상의 관찰을 통해 세포 해부학적 관찰을 통해 측정할 수 없는 약물의 생체내 독성을 관측하는 시도가 계속 진행되고 있고 일부는 이미 상업화되어 이용되고 있다. 비록 현 단계의 기술 수준은 세포 해부학적 방법의 보완적인 자료로 독성 유전체학의 자료들이 활용되고 있지만 조만간 독성 판정용 DNA chip 등 유전체 정보를 이용한 정확도 높은 독성 판정 방법이 개발 될 것으로 사료된다[20].

요 약

인간 유전체 지도가 발표 된 이후 신약 개발의 많은 부분에서 유전체 정보가 활용되고 있으며 비록 현재의 초기 단계에서는 여러 가지 실험적 장애 요인으로 인해 그 활용도가 그리 크지 않다. 하지만 대부분의 약물들이 유전자 혹은 단백질과 같은 생체 구성 물질들의 활성을 조절하는 기작을 가지고 있다는 점을 고려할 때 유전체에 대한 전반적인 이해는 빠르고 정확한 신약 개발을 위한 필수적인 선결 조건이라 볼 수 있다. 따라서 이러한 접근은 제약회사를 비롯하여 병원, 연구소 등 관련 기관들에서 계속해서 활발히 진행되고 있고 멀지 않은 미래에 유전체 연구는 효율적인 신약 개발의 핵심 기술로 자리 매김 할 것으로 예측된다

참고 문헌

1. <http://www.nyu.edu/classes/ytchang/book/n003.html>
2. Drews J. Drug discovery: a historical perspective. *Science* 287, 1960-1964 (2000)

3. Bailey D., Zanders E., Dean P. The end of the beginning for genomic medicine. *Nature Biotechnol.* **19**, 207-209 (2001)
4. Chicurel ME., Dalma-Weiszhausz DD. Microarrays in pharmacogenomics - advances and future promise. *Pharmacogenomics* **3(5)**, 589-601 (2002)
5. Nees M., Woodworth CD. Microarrays: spotlight on gene function and pharmacogenomics. *Curr Cancer Drug Targets* **1(2)**, 155-75 (2001)
6. Figeys D. Adapting arrays and lab-on-a-chip technology for proteomics. *Proteomics* **4**, 373-82 (2002)
7. Albala JS. Array-based proteomics: the latest chip challenge. *Expert Rev Mol Diagn* **2**, 145-52 (2001)
8. Narayanan, A., Wu, X., Yang R. Mining viral protease data to extract cleavage knowledge. *Bioinformatics* **18**, S5-S13 (2002)
9. Agrafiotis DK., Lobanov VS., Salemme FR. Combinatorial informatics in the post-genomics ERA. *Nat Rev Drug Discov* **5**, 337-46 (2002)
10. <http://www.isip.com/antisens.htm>
11. Zhang, H., Cook, J., Nickel, J., Yu, R., Stecker, K., Myers, K., Dean, N. M. Reduction of liver fas expression by an antisense oligonucleotide protects mice from fulminant hepatitis. *Nature Biotechnology* **18(8)**, 862-867 (2000)
12. Shanahan, W. ISIS 2302, an Antisense Inhibitor of Intercellular Adhesion Molecule 1 (ICAM-1). In: *Antisense Drug Technology: Principles, Strategies, and Applications* (ed. Crooke, S.T.), pp. 773-794. Marcel Dekker, Inc., New York (2001)
13. Ferentz AE. Integrating pharmacogenomics into drug development. *Pharmacogenomics* **4**, 453-67 (2002)
14. <http://snp.cshl.org/>
15. O' Shaughnessy J., Vukelja SJ., Marsland T., Kimmel G., Ratnam S., Pippen J. Phase II trial of gemcitabine plus trastuzumab in metastatic breast cancer patients previously treated with chemotherapy: preliminary results. *Clin Breast Cancer Suppl* **1**, 17-20 (2002)
16. Mazzucchelli R., Corvasce S., Violin M., Riva C., Bianchi R., Deho L., Velleca R., Cibella J., Bada M., Moroni M., Galli M., Balotta C. Role of CCR5, CCR2 and SDF-1 gene polymorphisms in a population of HIV-1 infected individuals. *J Biol Regul Homeost Agents* **15**, 265-71 (2001)
17. Seufferlein T., Boehm BO. The impact of pharmacogenomics on gastrointestinal cancer therapy. *Pharmacogenomics* **5**, 625-33 (2002)
18. Tribut O., Lessard Y., Reymann JM., Allain H., Bentue-Ferrer D.. Pharmacogenomics. *Med Sci Monit* **7**, RA152-63 (2002)
19. Johnson JA. Drug target pharmacogenomics: an overview. *Am J Pharmacogenomics* **4**, 271-81 (2001)
20. Waring JF., Halbert DN. The promise of toxicogenomics. *Curr Opin Mol Ther* **3**, 229-35 (2002)