

림상에서 프로테옴해석의 리용

장 현 성

위대한 령도자 김정일동지께서는 다음과 같이 교시하시였다.

《정보기술, 나노기술, 생물공학을 발전시켜야 새 재료, 새 에너르기, 우주기술, 핵기술과 같은 첨단과학기술분야와 기계, 금속, 채취공업, 경공업, 농업을 비롯한 응용기술분야를 획기적으로 발전시킬수 있으며 시대의 추세에 맞게 경제를 현대화, 정보화할수 있고 나라의 경제구조도 개변할수 있습니다.》(《김정일선집》 증보판 제22권 21페이지)

프로테옴해석은 세포나 조직에서 기능적으로 발현되는 모든 단백질의 구조와 기능을 포괄적으로 해석함으로써 인체안에서 진행되는 복잡한 생리적 및 생화학적과정들을 해석하며 최종적으로 분자수준에서 질병을 진단하고 예방치료하는데 그 목적이 있다.

프로테옴해석은 단백질에 대한 연구에서 보다 높은 단계로 되며 병없이 오래오래 살려는 사람들의 요구를 충족시키는데서 큰 역할을 하게 될것이다.

1. 프로테옴에 대한 일반적리해

유기체안의 단백질에 대한 전게놈수준에서의 해석은 이미 2차원젤전기영동의 출현을 전후한 시기인 20여년전부터 구상되어온것이다. 특히 일부 연구자들은 인체내 모든 단백질을 분리하고 목록화할 구상을 발표[1-3]하였다.

프로테옴은 유기체에서 만들어지는 모든 단백질을 서술하기 위한 용어로 제기되었다.[4] 게놈(genome)이라는 단어가 염색체(chromosOME)와 그 유전자(GENE)들의 총체라는 의미의 합성어라면 프로테옴(proteome)이라는 단어는 염색체(chromosOME)와 그에 의해 암호화된 단백질(PROTEin)들의 총체라는 의미의 합성어로서 게놈의 매개 유전자에 대응하는 모든 단백질을 말한다.[5]

게놈과 관련된 학문분야를 게놈학(genomics)이라고 하듯이 프로테옴과 관련된 학문분야를 프로테옴학(proteomics) 일명 프로테옴해석이라고도 한다.

게놈은 사람유전정보의 총체이지만 단백질을 만드는 설계도에 지나지 않는다. 사람의 몸안에는 기본적으로 동일한 게놈을 가진 여러가지 크기의 세포들이 있는데 세포에 따라 어떤 유전자가 활성화되는가 하는것이 다르고 만들어지는 단백질도 다르다. 하나의 유전자로부터 구조가 서로 다른 여러가지 단백질이 만들어진단. 단백질은 DNA로부터 전사된 mRNA를 주형으로 하여 합성되는데 단백질은 폴리펩티드로 합성된 다음 당분자와의 반응, 린산화 등을 거친 후 기능이 발현되거나 조절된다.

개별적유전자는 그 산물인 단백질을 통하여 여러가지 생물학적기능(효소와 호르몬, 전사조절인자, 여러가지 접수체와 막단백질, 저장단백질 등)을 나타낸다. 게놈정보에서 많은 DNA의 염기배열이 결정되고 유전에 대한 많은 정보를 장악할수 있었지만 염기배열정보가 많아지면서 DNA의 해석만으로는 밝힐수 없는것이 적지 않다는것을 알게 되었다. 그 하나가 유전자번역생성물인 단백질의 기능이다. 게놈정보를 효과적으로 리용하자면 실제로 정제한 단백질과 그것들을 암호화하는 게놈DNA와의 대응관계를 밝히며 단백질의 생화학적 및 물리화학적특성을 분석하고 모든 유전자의 번역생성물에 대하여 기능을 밝혀야 한다. 기

능을 나타내는 단백질의 움직임과 수식화정도, 단백질분자사이의 호상작용 등에 대한 해석과 염기배열정보를 대응시키는 방법으로 매개 유전자의 기능을 명백히 밝힐수 있다.

유전자의 번역생성물가운데서 자료기지를 리용한 상동단백질의 검색 등에 의하여 기능을 추정할수 있는 단백질의 비율은 전체의 30~50%정도이다. 이 비율은 고등한 생물일수록 더 작아진다.

프로테옴해석에서는 서로 다른 환경에서 모든 생육단계에 있는 기관에서 발현되는 단백질을 연구대상으로 하고있다. 최근에 분석방법과 분석계기 등이 급속히 발전됨으로써 프로테옴연구도 가까운 앞날에 완료되리라 보고있다. 프로테옴해석은 게놈해석을 기초로 한것이므로 게놈해석이 진척되지 않았던 1900년대이전에는 존재하지 않았으며 종래의 단백질해석만이 진행되었다. 인류는 유전자의 총개수를 정확히 알게 되고 인간유전자정보와 질병의 근원을 알게 되었으며 최종적으로 인간의 생존과 노화, 질병, 죽음 등 절대다수 유전정보의 게놈지도편람을 분석할수 있게 되었다. 그리하여 인간의 질병진단, 새로운 약물의 제조, 새로운 치료방법의 탐구 등 력사상 전례가 없는 일대 변혁이 일어나 생명과학과 생물기술의 급속한 발전을 가져올수 있게 되었다.

그러나 감염증을 제외한 많은 질병들은 대체로 유전적인 원인과 함께 생활환경의 변화와 같은것이 겹쳐서 발생한다. 전형적인 유전병에서는 원인유전자와 림상증상이 직접 연관되어있으므로 게놈정보로부터 병발생의 물림새를 대체로 추정할수 있지만 생활습관과 환경요인의 작용으로 발생하는 질병에서는 여러개의 유전자가 관계되고 이때 여러가지 단백질의 기능적인 평형이 파괴되어 복잡하게 얽히므로 이러한 질병의 병태를 정확히 파악할수 없다. 그러므로 이러한 병발생의 물림새를 밝히기 위해 여러개 단백질의 변화를 동시에 해석하는 문제가 제기된다.

이전의 단백질연구에서는 단백질을 순서대로, 개별적으로 해석하는 방법을 리용하였기때문에 여러가지 인자에 의하여 발생하는 질병을 해석할수 없었다. 그러나 프로테옴해석에서는 기능적으로 발현되는 전체 단백질을 모두 해석하여 린산화와 당사화 등 번역후 수식에 관한 정보를 얻으므로 여러가지 인자가 작용하는 질병의 병태해석을 진행할수 있다.[23, 24]

2. 림상에서 프로테옴해석의 일반적공정

림상에서 프로테옴해석은 최근시기 사람들이 커다란 관심을 가지고있는 미개척과학분야이다. 림상에서 프로테옴해석은 만성질병들이나 암 그리고 신경퇴행성병리과정들을 해명하려는 림상의사들이 질병의 분자적측면들을 특징지을수 있게 해준다.[6-9] 프로테옴해석의 중요하고도 세부적인 계획작성의 기초는 단계별실험시료들에 포함되어있는 단백질종류와 그 함량에 대한 정보이다. 그러나 프로테옴해석의 실제적인 목적은 단백질목록정보 그 자체가 아니라 림상실천에서 응용할수 있는 적실한 자료를 얻는것이다. 다시말하여 실제로 분자적특징을 밝히고 영향력있는 경로들을 찾으며 가능한것 진단, 치료결과에 대한 예측에 도움이 되는 알려진 생물표식자들을 확정하고 질병의 물림새들을 밝히는것이다.[10-13] 이것은 이미 많은 연구자들이 제기한바와 같이 기초적인 프로테옴학이 림상연구에 전 환되고있다는것을 의미한다.[14]

프로테옴해석의 연구대상인 실험시료속의 단백질은 외적요인의 영향으로 변화될수 있으므로 그 보관처리와 실험조작과정에 주의를 돌려야 한다.

외적요인으로는 생물학적차이, 분석전차이, 분석중차이 등이 있다. 생물학적차이에는 개체들사이의 차이와 개체내에서의 차이가 있다. 분석전차이는 시료처리과정에서의 오차

고 분석중차이는 실험기구와 설비에 의한 오차이다. 이러한 차이들은 병과는 련관이 없으나 프로테옴에 대한 분석결과해석에서 복잡성을 조성할수 있다. 단백질의 변화가 병의 발생발전과정에 생긴것인가를 알아내기 위하여서는 실험설계를 잘하는것과 함께 좋은 설비를 사용하고 정확한 통계학적방법을 리용하여야 한다. 또한 자동기계장치를 리용하여 실험오차를 없애야 한다. 시료는 많은 량을 수집하여야 하고 각종 질병들에 대하여 모두 수집하여야 한다. 그래야 질병의 전반상태를 정확히 알수 있다.

실험시료의 세포류형과 조직구조, 단백질조성성분은 매우 복잡하다. 실험에 앞서 시료를 합리적으로 전처리하면 시료준비의 복잡성을 없애고 실험분석에 도움을 줄수 있다.

실험시료에는 조직시료와 액체(혈액 등)시료가 있는데 조직시료에 대하여서는 레이자포획미세절단기술(LCM)을 적용한다. 이 기술은 현미경과 레이자빔을 리용하여 특정한 류형의 세포를 절단분리하는것이다. 이 방법으로 프로테옴해석을 위한 조직시료의 순도를 보장할수 있다. 혈액시료는 정보량이 매우 풍부하고 단백질조성성분이 복잡하며 단백질들의 함량도 매우 넓은 범위에서 변한다. 혈장에는 함량순위가 22번째까지인 중요 단백질이 99%를 차지하고 나머지 1%는 5 000종이상의 단백질, 펩티드가 차지한다. 그 가운데서 량적으로 매우 적어 그 존재를 확인하기가 힘든 단백질과 펩티드가 임의의 질병관련표식자로 될수 있다는데로부터 프로테옴해석이 심화되고있다. 복잡한 혈액단백질 가운데서 함량이 적은 단백질의 표식자를 찾으려면 알부민과 면역글로블린을 비롯하여 함량이 높은 단백질들을 가능한대로 제거하여야 한다. 이때 함량이 낮은 단백질이 함께 제거되지 않도록 하여야 한다.

프로테옴해석기술은 단백질분리기술과 단백질동정기술, 프로테옴정보학으로 구성되어 있다.

단백질분리기술로는 2차원겔전기영동(Two-Dimensional Gel Electrophoresis)이 리용되고 있다.

단백질동정기술로는 질량분석법이 기본방법으로 되고있다.

1990년대에 질량스펙트르분석(MS)은 림상화학분야에서 획기적인 비약이 일어나게 하였으며 이때부터 약물과 스테로이드성물질들을 포함하여 작은 분자들을 분석하는 표준정량방법으로 되고있다.[15] 진단에 도움이 되는 단백질들과 펩티드들을 리용하는 MS의 림상적역할이 더욱더 커지고있다.[16] 현재 액체크로마토그래프-직결식질량분석(LC-MS/MS)에 기초한 프로테옴해석으로 조직에 있는 1만개이상의 단백질들을 검출할수 있[17]을뿐아니라 세포질에 있는 몇천개의 단백질들도 검출할수 있다.[18] 그러나 현재까지도 여러가지 기술적 및 분석에서의 난점들이 정량프로테옴의 주요림상적발전을 저해하고있다. 한가지 실례로 세포질프로테옴에 대한 MS의 분석검출한계가 대체로 1~10ng/mL이상으로서 진단에 리용되는 세포질단백질들과 펩티드들의 생리적농도범위보다 높은것[19]이다. 림상에서 프로테옴해석의 목적은 비용은 들지만 명백한 단백질이성체들을 포함하여 가능성있는 수백가지의 진단성단백질들의 효과적인 다중화를 실현하자는것이다. 일부 세부적인 분석계획들이 특별히 주목을 끄는데 그것은 이 계획들이 수많은 림상실험실들에서 이미 진행되고 있는 거친 질량분석의 가동환경에 맞기때문이다.[20]

프로테옴정보학은 프로테옴해석기술의 중요한 기초이다. 프로테옴정보학에서는 분리, 동정을 거쳐 얻은 실험결과를 컴퓨터에 입력시켜 가공, 분석하고 최종적으로 생물학적인의미를 가지는 정보를 얻어낸다.

이와 같이 2차원겔전기영동과 LC-MS/MS 등과 같은 감도와 정밀도가 높은 대규모분석기술을 리용하여 생체조직과 혈액속에 존재하는 생물표식자 및 약물표적후보단백질들을

검출, 동정한다. 웨스턴 블롯팅법, MS/MS, RNA 간섭법, 유전자과피법과 같은 여러가지 방법과 수단들을 리용하여 동정된 질병관련단백질의 특징, 기능, 질병과의 연관 등을 해명한다. 또한 효소결합면역흡착검정법(ELISA)과 MS/MS로 여러 시료에서 질병관련단백질들이 특이적으로 발현되는가를 분석한다. 프로테오믹정보학을 통하여 생물학적인의의를 가지는 정보를 분석한다.

3. 암성질병의 조기진단 및 치료결과예측에서 생물표식자의 응용전망

일반적으로 종양은 유전성질병으로 알려져있다. 연구자들은 일련의 암유전자들의 활성화와 암유전자의 발현을 억제하는 유전자들의 사멸이 암을 발생시킨다고 보고있다. 그러나 유전자산물의 측면에서 볼 때 종양은 프로테오믹해석의 질병대상으로 볼수 있다. 즉 한 조의 단백질들에서 일어난 변화는 전체 신호경로에서의 이상을 발생시키고 최종적으로는 암을 발생시킨다. 그러므로 종양은 림상에서 프로테오믹해석의 좋은 질병모형으로 되며 연구 대상으로 된다. 그중에서도 암을 조기에 발견할수 있는 생물표식자를 찾아내고 새로운 치료약을 개발하는것이 매우 중요하다. 1개의 종양표식물은 대부분의 종양들에 대하여 정확한 진단근거로 될수 없다. 여러개의 종양표식자를 호상연관속에서 관찰하고 이것으로 종양의 전체적인 특징을 반영하는것은 단일종양표식물만 리용하는것보다 우월하다.

한 연구자는 3종의 란소암표식자를 발견하였는데 이것들을 각각 단일종양생물표식자로 리용할 때에는 란소암감별능력이 암항원125(Cancer Antigen 125: CA125)보다 높지 못하였다. 그러나 그것들을 CA125와 함께 리용하면 감별능력을 훨씬 높일수 있었다. 그중에서 2종의 란소암표식자를 CA125와 함께 리용할 때 감수성은 94%였고 CA125를 단독으로 리용할 때에는 81%였다.

프로테오믹해석방법을 리용한 생물표식자 및 약물표적단백질의 탐색, 그것의 평가 및 리용방법에 대하여 란소암연구를 실례들어 보기로 한다.

고감도대규모정량방법을 리용하여 란소암조직과 배양세포속의 단백질들을 분석하였다. 결과 란소암에 의하여 발현량이 유의하게 변화되는 약 40개의 단백질들을 검출, 동정할수 있었다. 동정된 란소암관련단백질가운데서 아넥신 IV와 질병과의 연관을 검증하는 실험을 진행하였다. 먼저 웨스턴 블롯팅법을 리용하여 암조직과 배양세포에서 이 단백질의 발현량을 조사하였다. 그 결과 아넥신 IV는 란소암세포에서 발현량이 특이적으로 증가된다는것이 확인되었으며 이 단백질의 발현이 전사수준에서 조절된다는것이 밝혀졌다. 한편 아넥신 IV와 함께 검출, 동정된 대부분의 란소암관련단백질들의 발현도 전사수준에서 조절된다는것을 알수 있었다. 이 유전자의 발현을 siRNA로 억제하였을 때 단백질의 합성과 암세포의 증식도 억제되었다. 실험결과로부터 연구자들은 아넥신 IV를 진단표식자 혹은 약물표적단백질로 리용할수 있다고 보았다.

면역침강법과 질량분석법을 함께 리용하여 란소암에 의하여 린산화가 향진 혹은 억제되면서 생긴 많은 단백질들을 검출, 동정하였다. 그가운데는 Stat3과 같이 린산화의 도움으로 유전자전사를 활성화함으로써 암세포증식을 촉진하는 단백질도 알려졌다. 란소암에서 린산화와 같은 번역후 수식이 단백질의 발현조절과 연관되어있을수 있다는것이 연구자들의 견해이다.

암조직 및 배양세포에서 검출, 동정된 란소암관련단백질들을 혈장에서 검출할수만 있다면 진단표식자로서의 리용가치가 클것이다. 이로부터 혈장에서 우와 같이 암과 관련된 단백질들을 검출할수 있는가를 조사하였다. 그러나 혈장에서 단백질의 양적변화가 매우 심

하므로 혈장단백질을 직접 MS/MS로 측정하면 알부민과 같이 많은 량으로 존재하는 단백질만 검출되고 진단표식자를 포함한 미량단백질들은 검출되지 않았다. 따라서 연구자들은 미량단백질분획을 분석하자면 미리 혈장에 많은 량으로 존재하는 단백질들을 제거할 필요가 있다고 보았다. 이로부터 중공섬유막 혹은 항체탐을 리용하여 혈장에 있는 일련의 단백질들을 제거한 다음 질량분석장치로 미량단백질들을 검출, 동정해보았다. 이 방법으로 3일 동안에 수백 μL 의 혈액에서 거의 3 000개에 달하는 단백질들을 검출, 동정할수 있었다. 또한 란소암조직 및 배양세포에서 검출된 대부분의 암과 관련된 단백질들이 질병에 따라 변동된다는것을 검출할수 있었다. 따라서 혈장속에서 질병과 관련된 단백질을 검출하기 위하여서는 새로운 검출기술을 개발하여야 한다는것이 명백해졌다.

다른 실례로 간세포암에 대한 연구를 보기로 한다.

간세포암발생의 주요원인의 하나인 C형간염비루스(HCV)의 감염으로부터 암으로 이행하는 초기단계의 병리학적물림새는 현재 해명중에 있다. 지금까지 간세포암조직속에서 생물표식자로 리용할수 있는 많은 단백질들이 발견되었다.[21, 22]

최근 연구자들은 간세포암조직과 정상조직에 대한 프로테옴해석을 하고 C형간염과 관련된 암세포의 새로운 진단과 치료에 응용할수 있는 생물표식자들을 동정하는 실험을 진행하였다. 이 실험은 2차원겔전기영동과 액체크로마토그래프-직결식질량분석법(LC-MS/MS)으로 하였다.

암조직과 정상조직의 단백질에 대해서는 pH가 4~7인 띠형겔(11cm)에서의 등전점전기영동과 SDS-PAGE에 의한 2차원겔전기영동을 진행하였다. 전기영동이 끝난 다음 겔을 염색하였다. 암조직 및 정상조직단백질반점의 위치와 단백질발현량을 프로그램으로 해석하였다. 그리고 정상조직에 비하여 단백질반점의 염색세기가 1.5배이상($p < 0.05$) 증가 혹은 감소된 암조직의 단백질반점을 구체적으로 분석하였다. 통계적인 해석을 한 다음 염색세기가 증가(10개) 혹은 감소(12개)된 22개의 암조직단백질반점을 관찰할수 있었다. 이것들을 염색제로 염색하고 겔로부터 분리한 다음 LC-MS/MS로 분석하였다. 결과 C형간염비루스에 감염된 간조직에서 22개의 단백질들이 특이적으로 발현된다는것을 발견하였다. 정상조직에 비하여 암조직에서 22개중 10개 단백질들의 발현수준은 높았지만 나머지 12개는 낮았다. 이 단백질들을 간세포암에 대한 진단 및 치료에서 새로운 생물표식자로 리용할수 있겠는가 하는것은 더 확인해보아야 한다.

또 다른 실례로서 외과적으로 절제술을 진행한 취장암환자 20례에서 수술전과 수술후 2주일에 혈청단백질과 펩티드에 대한 프로테옴해석을 한 결과 총 85개의 분획가운데서 수술전과 수술후변화가 제일 큰것이 6.6kD와 6.4kD인 펩티드였는데 6.6kD 펩티드는 아포지단백 C1, 6.4kD는 그 N말단에서 2개의 아미노산이 결손된 펩티드라는것이 밝혀졌다. 그리하여 아포지단백 C1과 그 대사산물이 취장암환자의 수술후 예후를 반영하는 지표로 될수 있다는것이 밝혀졌다.

맺 는 말

프로테옴해석에서는 기능적으로 발현되는 전체 단백질들을 모두 해석하여 린산화, 당사슬화 등 번역후 수식에 관한 정보를 얻음으로써 여러가지 유전자가 관계되고 단백질의 기능적인 평형이 파괴되어 복잡하게 얽힌 질병들의 병태해석을 진행할수 있다.

참 고 문 헌

- [1] N. Anderson; Nature, **278**, 122, 1979.
- [2] N. Anderson et al.; Clin. Chem., **28**, 739, 1982.
- [3] N. Wade; Science, **211**, 33, 1981.
- [4] D. Swinbanks; Nature, **318**, 653, 1995.
- [5] Sandor Suhai; Genomics and Proteomics: Functional and Computational Aspects, Kluwer Academic Publishers, 107~120, 2000.
- [6] L. Beretta; Nat. Methods, **4**, 785, 2007.
- [7] H. J. Issaq et. Al.; J. Sep. Sci., **24**, 3484, 2011.
- [8] B. Grötrrup et al.; Proteomics, **22**, 4279, 2011.
- [9] A. C. Kroksveen et al.; J. Proteomics, **74**, 371, 2011.
- [10] R. Becker et al.; J. Cardiovasc Transl. Res., **4**, 231, 2011.
- [11] L. D. Shoemaker et al.; Neurosurgery, **70**, 518, 2012.
- [12] A. Plymoth et al.; Curr. Opin. Oncol., **23**, 77, 2011.
- [13] S. Camerini et al.; J. Chromatography, A **1381**, 1, 2015.
- [14] E. F. Petricoin et al.; Nat. Rev. Drug Discov., **9**, 683, 2002.
- [15] S. K. Grebe et al.; Clin. Biochem. Rev., **32**, 5, 2011.
- [16] A. N. Hoofnagle et al.; J. Immunol. Methods, **347**, 3, 2009.
- [17] J. R. Wiśniewski et al.; Clin. Appl., **7**, 225, 2013.
- [18] P. E. Geyer et al.; Cell Syst., **2**, 185, 2016.
- [19] G. L. Hortin et al.; J. Proteome, **73**, 629, 2010.
- [20] J. Albrethsen et al.; Clinica Chimica Acta, **477**, 18, 2018.
- [21] 大場雄介; 生化学, **76**, 1, 16, 2004.
- [22] 相場英樹; 生化学, **78**, 4, 321, 2006.
- [23] 小西博昭; 生化学, **79**, 8, 781, 2007.
- [24] 宮木賢一; 生化学, **78**, 12, 1131, 2006.

주체109(2020)년 4월 5일 원고접수

Applications of Proteomics Analysis in the Clinical Practice

Jang Hyon Song

In proteomics analysis, we should investigate all proteins expressed functionally and get information about post-translational modifications such as phosphorylation and glycosylation, so that pathophysiological analysis of diseases which are related with several genes, broke down in functional equilibrium of proteins and get entangled complicatedly can be progressed.

Keywords: proteome, analysis, clinical practice