

바이오연료 연구와 Metabolomics

박정진

2000년 휴먼 지놈이 완벽하게 해독되었으나 생명현상을 설명하기에는 단백질을 코딩하는 유전자의 수가 너무 적다는 것이 밝혀짐에 따라 휴먼 지놈 해석 이후의 생명현상을 이해하기 위한 functional genomics의 연구가 중요한 이슈로 부각되고 있다. Functional genomics의 가장 궁극적인 목적은 바로 유전자와 표현형과의 상관관계를 정확히 밝히는 것이다. 생명체의 표현형은 유전자와 환경의 상호작용의 결과로 나타나며 이러한 표현형은 주로 생명체가 항상성을 유지하려는 세포내 소기관의 대사산물에 의해 조절된다. 그러므로 대사과정에 있는 대사산물들을 분석하는 것과 그 대사산물의 역할을 알아보는 metabolomics는 functional genomic era에서 중요한 역할을 담당하게 된다.

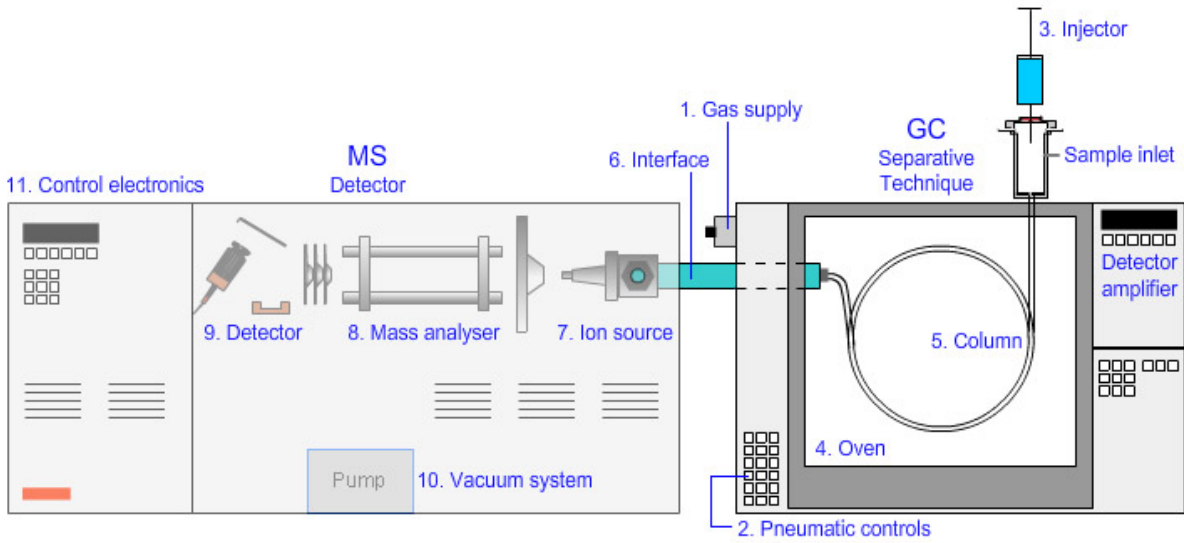
대사산물과 관련된 실험 단계로는 일반적으로 다음의 네 가지 단계로 분류할 수 있다. 첫번째로 특정 유전자 또는 단백질과 직접적으로 연관이 있는 물질만을 대상으로 분석하는 'target analysis'가 있다. 이는 주로 screening을 목적으로 하는 실험에 주로 이용된다. 두번째로는 특정한 pathway 상에 존재하는 대사산물을 분석하는 'metabolite profiling' (또는 metabolic profiling이라고도 한다)이 있다. 이 방법은 주로 약품의 분해 경로 등을 연구할 때 사용된다. 세번째로는 특정 생명체 내의 모든 대사과정에 존재하는 대사산물을 다루는 'metabolomics'가 있다. 여기서 metabolomics란 이미 알려져 있는 물질뿐만 아니라 아직까지 밝혀지지 않은 대사산물의 동정도 포함한다. 하지만 이러한 방법은 개개의 모든 대사산물을 분석해야 하므로 샘플의 숫자가 많다거나 산업체나 병원에서 진단용으로 적용하기에는 적합하지 않다. 그러므로 이러한 경우에는 대사산물의 생산 기작에 따른 각각의 class 별로 파악하는 'metabolic fingerprinting'을 사용한다.

현재의 기술 수준으로는 어떠한 종류의 분석기기로 metabolome전체를 커버하는 profiling은 불가능하다. 분석기기의 종류에 따라 분석 가능한 metabolite의 종류가 한정되며 민감도에도 큰 차이가 있다. Functional genomics 입장에서는 다루어야 할 샘플의 수가 많으므로 분석기기의 처리 속도도 중요한 요소가 된다. 또한 metabolite를 정성적으로 파악하기 위해서는 두 가지 종류의 분석기기를 상호 연결하여 사용해야 할 경우가 많다. 일반적으로 식물 재료를 사용할 때는 HPLC를 사용하면 커버하는 범위가 넓고 정량적 데이터를 얻을 수 있으며 mass spectrometer(MS)와 연결하여 사용하면 authentic sample이 없어도 다양한 metabolite에 대한 정성분석이 가능해진다. 그러나 HPLC는 재

현성이 떨어지는 단점이 있다. 이에 비해 gas chromatography(GC)는 HPLC가 커버하기 어려운 지방산 계열의 화합물을 대상으로 할 때 매우 유리하고 데이터의 재현성과 안정성이 HPLC에 비해 월등히 높다. 또한 HPLC와 마찬가지로 MS와 연결하여 사용하는 것이 일반적이다. 다만 GC를 사용할 때는 샘플을 유도체화 시켜야하는 불편함이 있다. NMR 중 특히 $^1\text{H-NMR}$ 은 LC나 GC와 연결하여 사용하기도 하지만 단독으로도 정량성이 뛰어나고 상당한 정도의 정성 분석이 가능하며 샘플 전처리에 어려움이 없고 속도가 빠르다는 이점이 있다. 그러나 NMR은 민감도가 매우 낮다. FTIR은 정성분석적 정보를 별로 제공하지 못하지만 샘플의 전처리, 속도, 민감도 등에서 발군의 성적을 보여준다. 특히 최근에는 한꺼번에 384개의 샘플을 처리할 수 있도록 고안된 기종도 출시 되었다. $^1\text{H-NMR}$ 과 FTIR은 특히 metabolic fingerprinting에서 주로 사용되는 분석기기이다. 최근에 MS를 중심으로 괄목할 만한 기술적 진보가 이루어지고 있다. 이에 따라 capillary electrophoresis mass spectroscopy(CEMS)가 metabolomics에 광범위하게 사용되고 있는데 이는 비용, 선택성 (selectivity), 민감도, 정량정성 분석의 정도, 속도, 분석기기의 가격 등 다양한 결정 요소를 적절하게 타협적으로 만족시키기 때문이다. FTMS는 단번에 수백 혹은 일천 종 이상의 metabolite의 정량정성 분석이 가능하다는 이점 때문에 주목을 받는 분석기기이나 현재로서는 고가이므로 일반 연구자들이 사용하기 어려운 것이 현실이다.

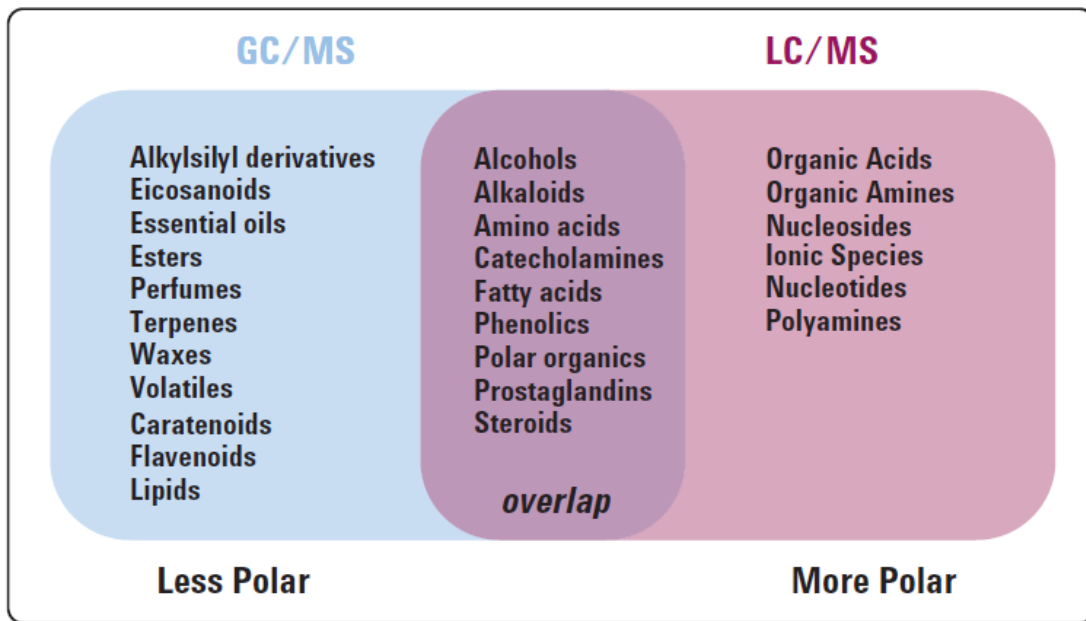
최근 다양한 종류의 대사산물 분석에 관한 필요가 늘어감에 따라, 보다 광범위한 데이터 확보를 위해서는 GC-MS와 LC-MS 모두를 필요로 한다. 하지만 실제로는 이렇게 하더라도 충분한 데이터 확보에 실패하거나, 다양한 변수들, 예를 들면 분석하려는 대사 산물의 화학적 성질, 샘플 매트릭스, 그리고 분석 단가에 따라 GC-MS나 LC-MS 둘 중에 하나를 선택해야만 하는 경우가 많다. 이번에는 이렇게 선택을 해야만 하는 경우, 보편적으로 어떠한 것들을 따져야 하는 지에 대해 알아보려고 한다. 그리고 GC-MS나 LC-MS를 이용한 일반적인 metabolomics 연구 방법 또한 살펴보도록 하겠다.

우선 가스 크로마토그래피와 질량분석기(GC-MS)는 휘발성 물질을 분석하는 데에 매우 효과적인 방법이다. 가스 크로마토그래피(GC)는 시료를 증발시켜 크로마토그래피관에 주입하고 비활성 기체 이동상의 흐름을 이용하여 용리시킨다. 대부분의 다른 형태의 크로마토그래피법과는 대조적으로 이동상은 분석물의 분자와 반응하지 않으며 충전제를 통하여 이들 분자들을 이동시키는 기능을 할 뿐이다. 그러므로 기본적으로 분석 물질은 휘발성이던지 아니면 화학적으로 유도체화 과정을 거쳐 휘발이 되도록 해야 한다.



<그림 1> 일반적인 GC-MS 개략도

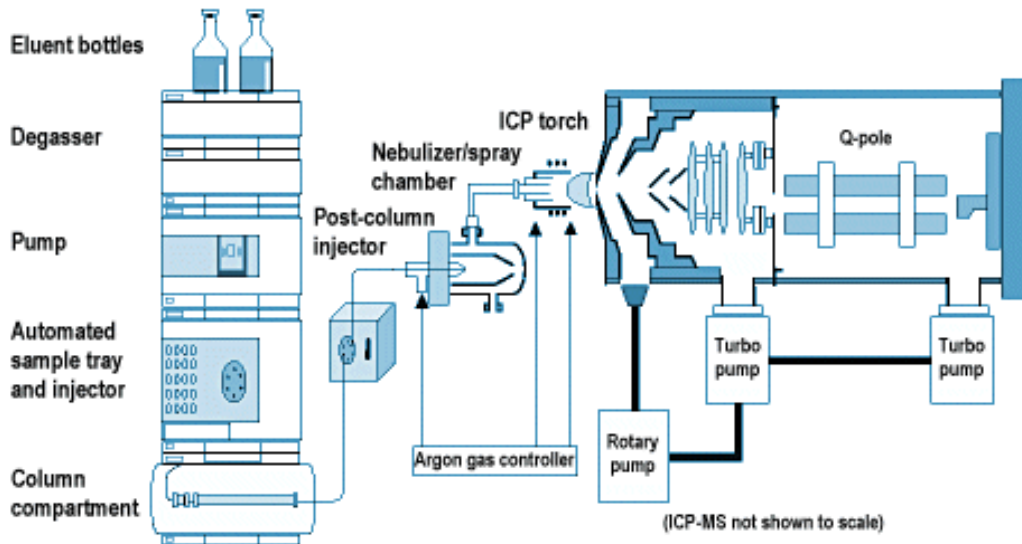
<그림 2>에서와 같이 GC-MS에 잘 들어 맞는 물질 들이 있으며, 이들 중 식물성 terpene이나 essential oil과 같은 것은 휘발성이면서 LC-MS에서는 이온화가 잘 되지 않는 성질을 가지고 있다.



<그림 2> 분석 대상물의 특성에 따른 일반적인 분석 방법

Liquid chromatography의 경우 대사산물이 휘발성이 필요 없으며, 유도체화 과정이 필요하지 않기 때문에 보다 넓은 범위의 물질을 분석할 수 있다. 하지만 GC-MS 처럼 물질 동정을 위한 spectral 라이브러리가 존재하지 않으므로, NIST 라이브러리 검색에는 이용

이 불가능하지만, 분자화 이온(molecular ion)이 대부분 검출되므로 이러한 분자값을 이용한 DB 검색에는 유리하다. 요즘 나오는 정밀한 질량 분석기의 경우 소숫점 이하 네째 자리까지 분리가 가능하다. 이러한 특성을 가진 LC-MS는 아직 사전 지식이 없는 물질 분석에 적합하다. 하지만 분석에 들어가는 비용은 LC-MS grade를 이용해야 하므로 (4L에 보통 10~30만원선) GC-MS 보다 많이 들어간다.



<그림 3> 일반적인 LC-MS의 개략도

질량분석기를 통하여 수집된 데이터는 다양한 구조의 저분자 대사체를 포함하고 있기 때문에 실제로 정성분석이 매우 곤란하여 실제로는 패턴분석을 통하여 분류하는 것이 일반적이다. 따라서 광범위한 데이터 안에 존재하는 대사체를 정성분석할 필요가 발생하였으며 이를 위하여 high resolution mass spectrometry가 사용되고 있다. High resolution mass spectrometry를 이용하여 parent의 분자량을 측정하고 MS/MS 기법을 사용하여 daughter ion을 비교하는 정성분석 방법이 시도되고 있으나 low resolution mass spectrometry의 경우 측정된 대사체가 0.5 m/z의 분자량 오차를 보인다면 DB에 너무나 많은 예상 후보 대사체가 있으므로 정성분석이 어려워진다. 따라서 high resolution mass spectrometry를 이용하여 정확한 분자량을 측정한 후 MS/MS fragments를 이용한 정성분석 시스템이 요구되고 있다. 예를 들어 MS의 resolution이 1ppm일 경우 이론적인 분자량이 100.00000인 물질은 99.9999~100.0001의 실험측정 분자량이 나오기에 그 범위의 monoisotopic 분자량을 가지고 있는 데이터베이스내의 예상 후보물질의 수는 10 ppm (99.999 ~ 100.001)에 비해서 현격히 줄어들게 되기에 구조를 구명하는데 매우 유용하다. 따라서 FT-ICR MS, Q-TOF MS, Orbitrap MS 같은 high resolution을 구현할 수 있는

질량분석기를 사용하는 연구가 많이 이루어지고 있다.

대사산물(metabolite)는 유전자 발현의 최종산물이기에 DNA, RNA, 또는 protein에 비해서 변화의 폭이 커서 유의성 있는 변화 관찰에 유리하다. 또한 다른 omics 연구 방법에 비해 경제적이라는 장점을 가지고 있다. 하지만 검출 가능한 물질이 아직까지는 수백개 수준으로 한계가 있으며 발현량 증폭이 불가능하다는 단점이 있다.

국내의 경우 지금까지 genomics 및 proteomics에 거의 대다수의 연구가 이루어져 왔으며 극소수의 기관에서 세포내의 metabolite profiling에 관한 연구를 수행해 왔다. 하지만 국내의 경우 아직 metabolomics와 proteomics의 통합 연구가 초기 단계이기 때문에 해외연수를 통해 선진 기술의 습득을 통한 인프라 형성이 요구된다.

Metabolomics 및 proteomics는 신약, 신규 항체, 진단 시스템 등 다양한 바이오산업과 연계되어 있어 향후 산업적 파급효과가 큰 연구 분야로 선진 각국에서는 각 요소 기술을 선점하기 위한 노력을 하고 있다. 그러므로 향후 국가 기술 경쟁력 확보를 위해서는 현재 이 분야를 선도하고 있는 선진국에 신진 인력을 파견하여 이 분야의 첨단 기술정보 및 연구 동향을 파악하고 상대국과의 제반 기술 교류가 절실히 필요할 것으로 보인다.