

GPCR 분류 및 기능 연구에 있어서 생물정보 데이터 분석 동향

허철구(한국생명공학연구원)

GPCR은 세포 표면에 가장 다양하게 존재하는 막단백질(transmembrane protein)로, 시중에 판매중인 치료제(therapeutic drug) 중 60% 이상이 이 막단백질을 타겟으로 한다. 그리고 GPCR의 구조적인 특징과 class 및 subfamily 등은 기능별로 잘 알려져 있다. 과거 GPCR을 찾아내는 연구 중에

가장 기본이 되는 일이 주어진 단백질 서열로부터 GPCR를 분류하는 일을 많이 수행하여 왔는데 이미 찾아진 GPCR들을 가지고 수학적인 모델을 이용하여 보다 쉽게 분류하는 연구를 진행하여 왔다. 예를 들면 수학적인 이론을 근거로 기계학습의 한 분야인 HMM 이론과 SVM이론(Manoj Bhasin *et al.*, 2005, N.A.R 그

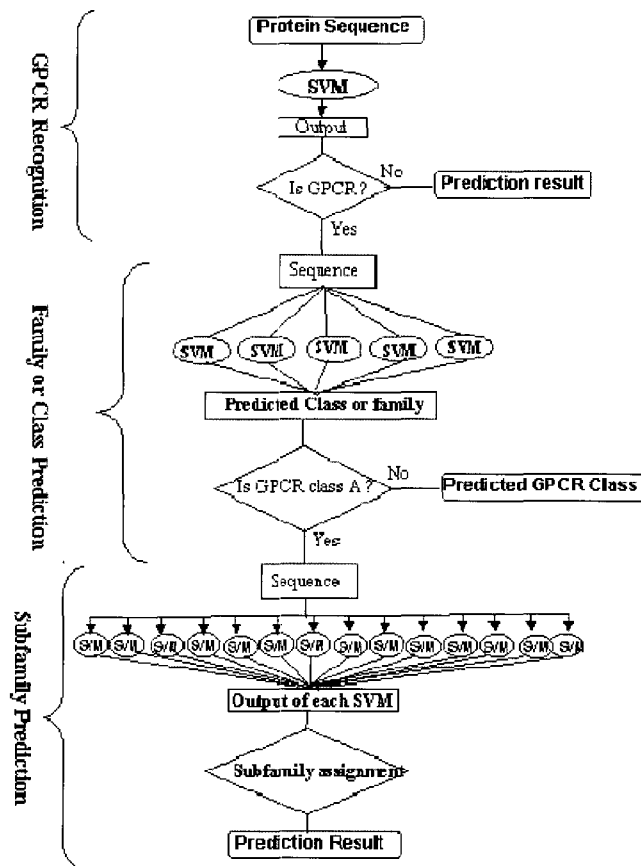


그림 1. SVM을 이용한 GPCR class, subfamily 분류

림 1)을 도입하여 GPCR를 분류하는 연구에서 HMM, SVM이론에 physicochemical property를 이용한 분류 기법등을 도입하면서 새로운 GPCR을 찾아내고 분류하는 연구(Maris *et al.*, 2006, Protein Science)는 계속 진행되고 있다.

또 2000년 초반부터 급격히 연구가 진행되고 있는 Human, Mouse Genome연구 분야 중 mRNA alternative splicing(AS)는 연구 결과가 활발히 진행됨에 따라서 새로운 GPCR서열을 찾아내고 분석하는 연구가 매우 많은 진전을 나타내고 있다. 하지만 여전히 단백질 기능에 있어서 mRNA의 AS 이후 효과에 많은 의문점을 가지고 있고 그 mRNA AS 이후 막단백질과의 연관성을 찾아내는 연구와 더불어 GPCR alternative splicing variant 연구도 매우 중요한 연구 분야로 나타나고 있다(M.S Cline *et al.*, 2004, PSB 그림 2). GPCR의 다양성은 주로 서로 다른 유전자에서 발현되는 GPCR에 의해서 증가되지만, splicing events에 의해서도 증가

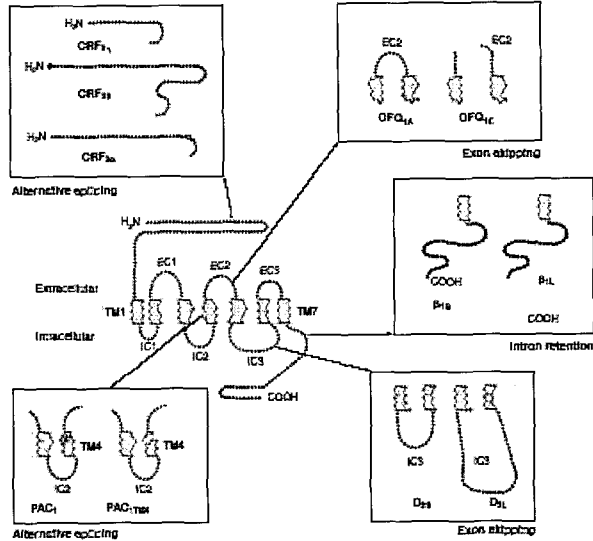


그림 2. GPCR alternative splicing

된다. (Kilpatrick *et al.*, 1999, Minneman *et al.*, 2001) 이 Splicing variants에 의해 생성된 GPCR isoform의 기능은 본래의 form과 다른 경우도 발견되고 있으며,

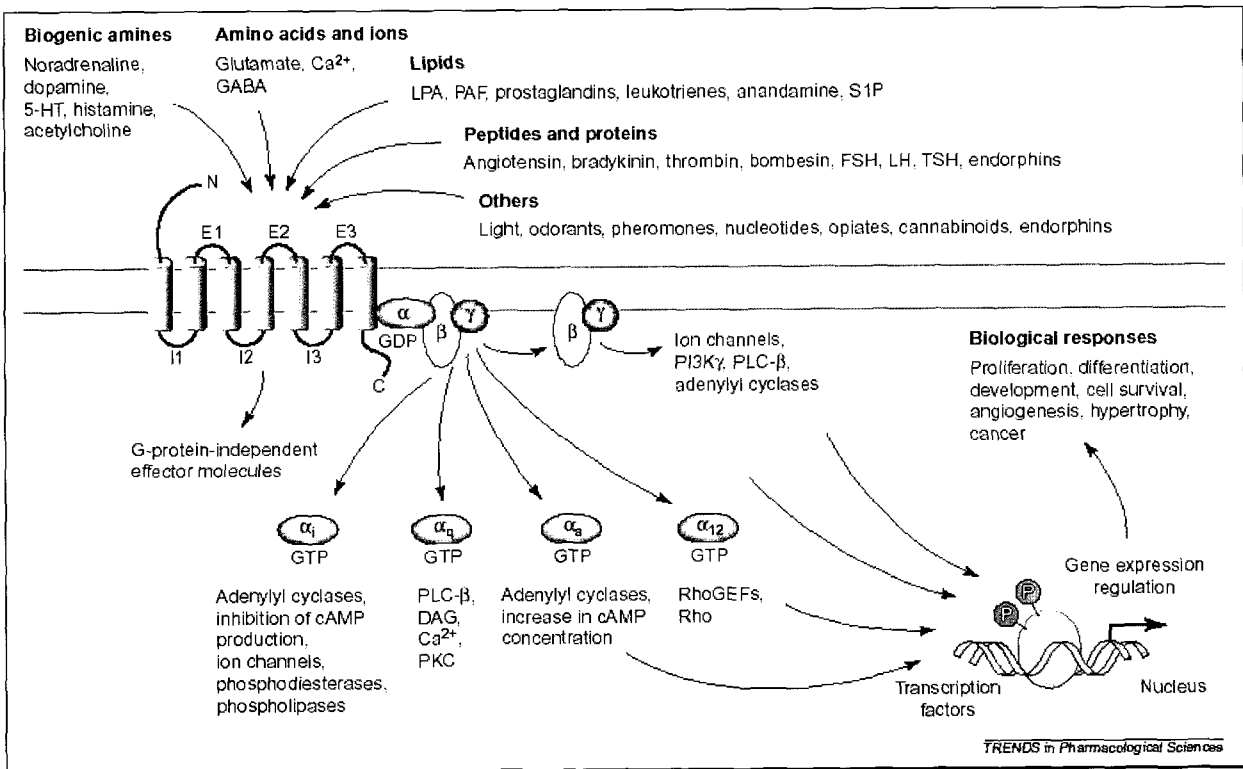


그림 3. GPCR의 생리적인 연구 모델



GPCR의 기능을 밝히고 분류하기 위해서는 GPCR isoform에 대한 연구도 최근 병행되고 있다. 생물정보 분야도 과거 각 분야별 연구 형태에서 실험실의 기술과 생물정보 관련 데이터가 늘어나면서 실험자들이 보다 복잡하고 정교한 정보를 요구하게 되고 있다.

따라서 생물정보 연구자들이 해결하여야 하는 문제들에 대한 복잡도가 높아지면서 생물학 지식과 전산학 기술, 수학, 통계학 분야 전문 지식이 상호 연결되는 형태의 네트워크형 연구 형태가 요구되고 있다. GPCR 서열을 정확하게 분류하고 GPCR AS에 대한 연구가 진행되면서 GPCR protein-protein interaction 개념을 도입한 GPCR과 G-protein간에 관련성 높은

연구 분야도 생물정보 연구에서 최근 많이 응용되고 있다. 이런 연구를 통하여 Orphan GPCR에 대한 이해도 역시 많이 달라지고 있는 것이다.

GPCR의 다양성과 생리적인 연구를 통하여 신약타겟을 발굴하는 분야도 실험적인 방법으로 연구하는 시대에서 복잡한 생물정보 분야와 병행하여 연구하는 방법론은 앞으로 생명공학 분야 연구에 있어서 새로운 전기가 될 것으로 기대된다. 결론적으로 유전체 관련 데이터의 증가는 생물정보 분야의 복잡한 분석능력을 요구하게 되고 실험 기반의 생명공학 분야 연구자들에게도 새로운 실험 모델을 적용하는 시대를 요구하고 있는지도 모른다.