

## 항원결정부위 예측을 위한 통합시스템 개발

### Development of Integration System for Epitope Prediction

진혜정, 이지후, 이인성, 김학용  
충북대학교

Hye Jeong Jin, Jihoo Lee, In Seoung Lee,  
Hak Yong Kim  
Chungbuk National University

#### 요약

질병 치료에 대한 패러다임이 과거 치료 위주에서 조기진단 및 예방의 개념으로 전환되고 있으며 진단용 마커와 단클론 항체 제작은 핵심 기술로 부각되고 있다. 현재까지의 항원결정부위(epitope) 예측은 단백질의 1차구조인 아미노산 배열 순서를 바탕으로 추출되어 진다. 하지만 항원결정부위는 수용성의 항체와 직접 결합하기 때문에 친수성 잔기가 차지하는 비율이 높아야 하며, 면역계가 쉽게 인지할 수 있도록 노출되어 있어야하고, 긴 선상의 폴리펩티드 단백질이 3차원 구조를 형성하기 위해 회전, 유연성 등이 요구된다. 따라서 한 가지 성질 중심으로 할 경우 오류가 나올 가능성이 있다. 이를 보완하기 위하여 본 연구에서는 친수성(hydrophilicity), 극성(polarity), 파묻힘성(buried residues), 접근성(accessibility), 회전성( $\beta$ -turns), 유연성(flexibility), 굴절성(refractivity) 등을 분석한 후 통합 예측시스템을 개발하였다. 이를 검증하기 위해 고양이 백혈병 바이러스의 항원결정부위를 예측해보았다.

## I. 서론

질병의 조기 진단과 함께 질병에 대한 예방이 중요하게 되면서 진단시장에 대한 관심이 고조되고 있다. 이에 따라 진단 키트제작의 기반인 타깃 단백질의 항원결정부위(epitope)를 예측하는 프로그램에 대한 개발도 활발히 이루어지고 있다.

현재 존재하고 있는 항원결정부위 예측 프로그램들은 아미노산 성질을 이용한 친수성(hydrophilicity), 극성(polarity), 파묻힘성(buried residues), 접근성(accessibility), 회전성( $\beta$ -turns), 유연성(flexibility), 굴절성(refractivity) 등 여러 인자들을 가지고 있다[1, 2]. 그러나 각 인자를 고려한 알고리즘에 따라 계산된 아미노산 경향성 수치가 다르기 때문에 분석의 차이가 나타난다. 또한, 한 가지 아미노산 성질만을 중심으로 항원결정부위를 예측할 경우 오류를 범할 수 있는 가능성이 있다.

본 연구에서는 항원결정부위 예측의 오류범위를 줄이고 특이성을 높이기 위해 여러 분석도구들을 통합하여 효율적이고 정확한 통합 예측시스템을 개발하였다.

## II. 본론

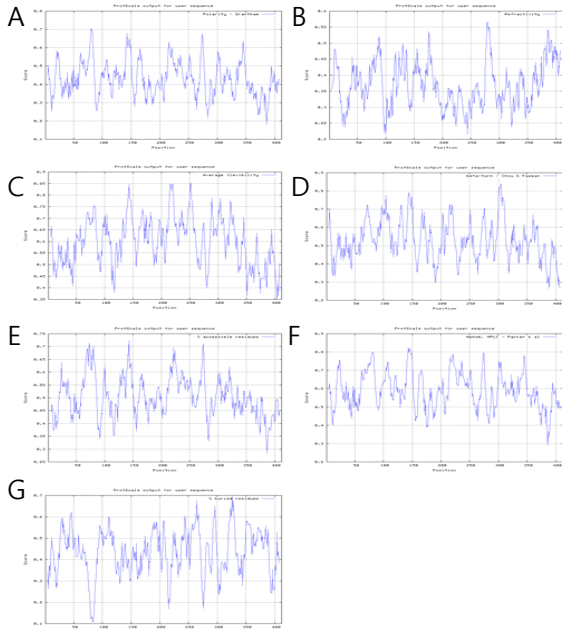
본 연구에서 아미노산 성질들은 ExPASy ProtScale을 이용하여 예측한 후 통합시스템을 구축하였다. 기존 예측프로그램인 IEDB Analysis Resource와 우리가 개발한 통합 시스템을 비교분석하기 위하여 고양이 백혈병 바이러스 gp70을 적용해보았다.

### 1. 통합시스템 구축

단백질 서열 중에서 면역 항체가 결합할 수 있는 항원부위를 예측하기 위해 ExPASy ProtScale에서 친수성, 극성, 파묻힘성, 접근성, 회전성, 유연성, 굴절성 등 아미노산 성질에 따라 각각의 항원결정부위를 분석하였다(그림 1). 예측 프로그램은 아미노산 경향성 수치를 알고리즘에 적용하여 항원부위를 선별한다[1, 2].

아미노산은 단백질의 구성요소로서 20종류가 있으며 각각의 아미노산은 서로 다른 구조와 성질을 가지고 있다. 또한 단백질 구성을 이루는 폴리펩티드의 성질은 아미노산의 종류와 위치에 따라 결정된다. 그렇기 때문에 프로그램에서의 측정된 아미노산 경향성 수치는 각각 파라미터에 따라 같은 아미노산이더라도 수치가 다르게 측정되어 분석의 종류와 목적에 따라 선별되어지는 항원결정부위에 차이가 나타난다.

아미노산 성질을 바탕으로 항원부위 분석하는 프로그램이 다양함에도 불구하고, 기존 항원부위 분석은 주로 친수성을 이용하여 예측하고 있다[3].



▶▶그림1. 아미노산 성질 분석

- A. 극성(polarity), B. 굴절성(refractivity),
- C. 유연성(flexibility), D. 회전성(β-turns),
- E. 접근성(accessibility),
- F. 친수성(hydrophilicity)
- G. 파묻힘성(buried residues)

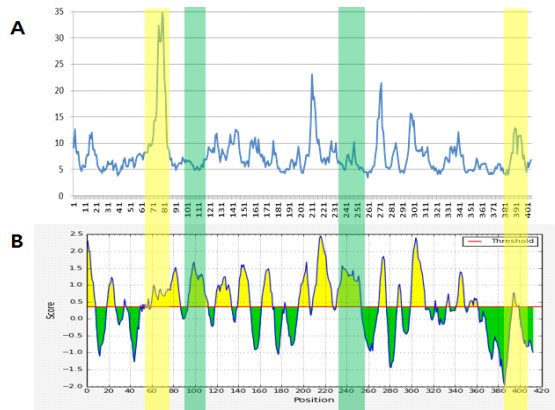
항체와 결합하는 항원인 에피토프는 혈액내의 수용액 상태이므로 친수성과 극성이 높아야하며 표면 노출성을 의미하는 접근성 또한 높아야 한다[2]. 단백질 안정성과 분자인식과정 역할을 하는 회전성도 중요하며 단백질의 3차 구조에서 외부로 노출된 부분은 굴절과 유연성이 있어야하기 때문에 우리는 아래와 같은 식을 이용하여 통합시스템을 구축하였다(그림 2A).

$$\text{항원부위 예측값} = \frac{(\text{친수성} + \text{극성} + \text{접근성} + \text{회전성} + \text{유연성} + \text{굴절성})}{\text{파묻힘성}}$$

## 2. 통합시스템과 기존 예측프로그램과의 비교

우리가 개발한 예측 통합시스템을 검증하기 위해 web 상에서 손쉽게 사용할 수 있는 항원부위 예측 프로그램인 IEDB Analysis Resource 결과와 비교/분석하였다. 친수성을 바탕으로 에피토프를 예측하는 IEDB Analysis Resource와 비교분석한 결과 비슷한 양상을 보였으나 이 예측 프로그램에서는 많은 서열부분이 항원부위로 선별되었다(그림 2B). 이러한 결과는 에피토프의 특이성이 떨어질 가능성이 있고, 이에 따른 오류의 확률이 높다고 해석할 수 있다. 반면에 우리가 개발한 예측 통합시스템을 이용하면 기존 프로그램의 무분별한 항원결정부위를

후보군선별을 보완함으로써 명확한 결과분석에 도움을 줄 수 있다.



▶▶ 그림2. 항원결정부위 예측결과

- A. 항원부위예측 통합시스템,
- B. IEDB Analysis Resource

## III. 결론

타깃 단백질의 항원결정부위를 정확하게 예측하여 고 성능의 진단키트를 제작하는 것은 매우 중요하다. 현재 일반적으로 사용하고 있는 항원결정부위 예측 프로그램은 아미노산 성질 중 하나의 특성(친수성)만을 고려하여 예측한다. 하지만 한 가지 아미노산 성질만을 중심으로 항원결정부위를 예측할 경우 오류를 범할 수 있는 가능성이 있다. 따라서 우리가 개발한 통합시스템은 항원결정부위를 정확성하고 효율적으로 예측할 수 있을 것이다. 향후 통합시스템을 이용하여 선별된 아미노산 부위가 진단키트에 실질적으로 이용할 수 있는 연구를 진행할 것이다.

## ■ 참고 문헌 ■

- [1] John M. Walker (ed): The Proteomics Protocols Handbook, Humana Press (2005), pp. 571-607
- [2] Jens Erik Pontoppidan Larsen, Ole Lund and Morten Nielsen, Improved method for predicting linear B-cell epitopes, Immunome Res, 2006; 2: 2.
- [3] Parker JM, Guo D, Hodges RS. New hydrophilicity scale derived from high-performance liquid chromatography peptide retention data: correlation of predicted surface residues with antigenicity and X-ray-derived accessible sites. Biochemistry. 1986 Sep 23; 25(19):5425-32.