

생물정보학기반 말라리아 항원부위의 예측

Prediction of epitope for Malaria based on bioinformatics

조시향, 김학용
충북대학교 생화학과

Si Hyang Jo, Hak Yong Kim
Chungbuk National University

요약

말라리아에 의한 감염 중 95% 이상을 차지하는 것은 열대열 원충과 삼일열 원충에 의한 감염이다. 각 원충들에 의한 감염은 초기증상이 유사하지만 각각의 치료방법이 다르게 적용되기 때문에 어떤 원충에 의한 발병인지 정확하고 빠른 진단이 필요하다. 본 연구에서는 생명정보학을 기반으로 각 원충에 특이적으로 발현되는 단백질 정보를 중심으로 epitope를 예측하고자 하였다. 먼저 NCBI에서 문헌조사 및 서열정보를 얻었으며, 두 원충에 대한 서열을 multiple alignment를 통해서 비교한 결과 90.2%의 유사성을 보였다. 확보된 특이 부위를 ProtScale Tool 프로그램을 이용하여 hydrophilicity를 분석하고 진단키트 제작에 사용할 수 있는 epitope를 예측하였다.

I. 서론

말라리아는 Plasmodium 속 원충이 적혈구와 간 세포 내에 기생함으로써 발병되는 급성 열성 감염증이다. 매년 약 3 ~ 5억의 말라리아 환자가 발생하며 그 중 100만 명 이상이 사망하고 있다. 사람에게 감염될 수 있는 말라리아 원충의 종류는 4종인데, 이 중 전 세계적으로 삼일열 원충과 열대열 원충이 95% 이상을 차지한다[1]. 열대열 원충에 의한 감염은 초기증상이 삼일열 원충에 의한 감염과 유사하지만 치료 방법이 다르다. 또한 열대열 원충에 의한 감염은 증상이 되면 황달, 혈액 응고 장애, 신부전, 간부전, 쇼크, 혼수 등의 급성 뇌증이 출현하기 때문에 신속한 치료를 위한 정확한 진단이 필요하다. 따라서 우리는 생명정보학을 기반으로 각 원충에 특이적으로 발현 되는 단백질 정보를 중심으로 epitope를 예측하고자 하였다.

II. 본론

1. 삼일열 원충과 열대열 원충의 L-lactate dehydrogenase 서열정보 추출

L-lactate dehydrogenase(LDH)는 열대열원충과 삼일열 원충에서 각각 발현되며 감염환자의 혈액에서 검출되어 현재 말라리아 진단 Pan 키트로 사용되고 있다. 따라서 각 원충별 LDH에 특이적인 epitope를 선정하기 위해 먼저 우리는 각 원충의 LDH 단백질 서열정보를 National Center for Biotechnology Information(NCBI)에서 얻었다.

2. 연구방법 및 연구결과

2-1. 단백질 서열 비교

삼일열 원충과 열대열 원충의 LDH의 유사성을 비교해보기 위해 단백질 서열 분석을 통해 단백질 간의 유사성을 보여주는 프로그램인 ClustalW2를 사용하여 multiple alignment를 수행한 결과 90.2%의 유사성을 보였다(그림 1).

SeqA	Name	Length	SeqB	Name	Length	Score
1	gi 66967948 gb AA559419.1	316	2	gi 76563843 gb ABA46355.1	316	90.19

CLUSTAL 2.1 multiple sequence alignment

```

gi|66967948|gb|AA559419.1|      MTFKPK|VLVSGSM|GGMATL|VQKNLGDV|MFQVKNMPOGH|ALDTSH 50
gi|76563843|gb|ABA46355.1|      MAFKAK|VLVSGSM|GGMATL|VQKNLGDV|LFDV|KNNMPPH|ALDTSH 50
*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*

gi|66967948|gb|AA559419.1|      SNMAYSNCK|VTGSNSVDLLK|GADV|VTAGFTKAPK|SKEINRDLLP 100
gi|76563843|gb|ABA46355.1|      TNMAYSNCK|VSGSNTVDLLK|GADV|VTAGFTKAPK|SKEINRDLLP 100
*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*

gi|66967948|gb|AA559419.1|      LNKI|MI|EGH|K|K|LCP|NAF|I|VVTN|PVD|MV|L|L|L|L|L|L|L|L 150
gi|76563843|gb|ABA46355.1|      LNKI|MI|EGH|K|K|K|C|P|NAF|I|VVTN|PVD|MV|L|L|L|L|L|L|L|L 150
*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*

gi|66967948|gb|AA559419.1|      GGVLDT|SRLK|Y|I|SQK|L|N|V|C|P|R|D|M|N|I|V|G|A|H|G|N|K|M|L|K|V|I|T|V|G|I|P|L 200
gi|76563843|gb|ABA46355.1|      GGVLDT|SRLK|Y|I|SQK|L|N|V|C|P|R|D|M|N|I|V|G|A|H|G|N|K|M|L|K|V|I|T|V|G|I|P|L 200
*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*

gi|66967948|gb|AA559419.1|      DEF|I|N|K|K|I|D|E|E|V|E|G|I|F|D|R|T|V|N|T|A|L|E|I|V|N|L|A|S|P|Y|A|P|A|A|I|I|E|A|E|S|Y 250
gi|76563843|gb|ABA46355.1|      DEF|I|N|K|K|I|D|E|A|E|L|E|A|I|F|D|R|T|V|N|T|A|L|E|I|V|N|L|A|S|P|Y|A|P|A|A|I|I|E|A|E|S|Y 250
*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*

gi|66967948|gb|AA559419.1|      LK|D|I|K|K|V|L|V|C|T|L|L|E|G|Q|V|G|H|S|N|I|F|G|G|T|P|L|V|I|G|G|T|G|V|E|V|I|E|L|Q|N|A|E|K|T 300
gi|76563843|gb|ABA46355.1|      LK|D|I|K|K|V|L|I|C|T|L|L|E|G|Q|V|G|H|S|D|I|F|G|G|T|P|V|L|G|A|N|G|V|E|V|I|E|L|Q|N|S|E|K|A 300
*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*

gi|66967948|gb|AA559419.1|      KFDEA|V|A|E|T|K|R|M|K|A|I| 316
gi|76563843|gb|ABA46355.1|      KFDEA|I|A|E|T|K|R|M|K|A|A| 316
*****
    
```

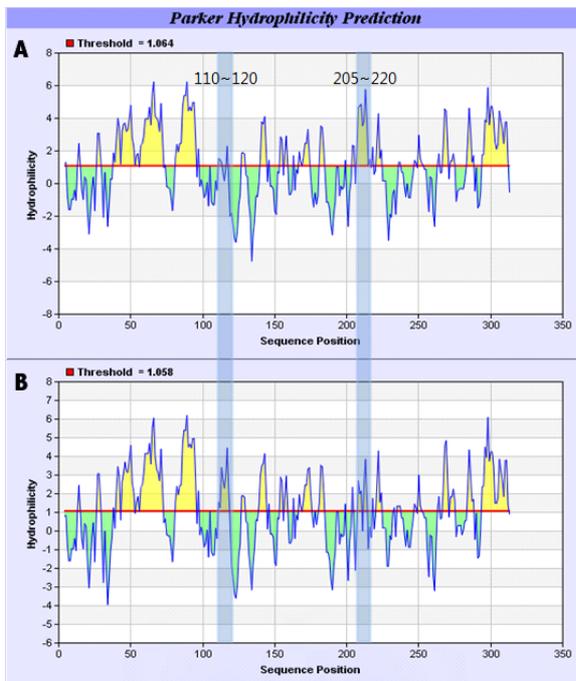
▶▶ 그림 1. 삼일열 및 열대열 원충의 서열비교

gi|66967948|gb|AA559419.1 : 삼일열 원충 LDH,
gi|76563843|gb|ABA46355.1 : 열대열 원충 LDH

*, fully conserved residue; ;, strongly similar properties;
, weakly similar properties

2-2. ProtScale Tool 프로그램을 이용한 hydrophilicity의 분석비교

우리는 삼일열 원충과 열대열 원충의 LDH를 Immune Epitope Database(IEDB)를 이용하여 각각에 특이적인 epitope를 예측해 보았다. 이 epitope를 실험적으로 사용하기 위해서는 무엇보다도 solubility가 중요하다. 따라서 우리는 단백질 아미노산 서열의 친수성과 소수성을 보여주는 hydrophilicity prediction을 통해 epitope의 친수성 여부를 확인하였다(그림2). 또한 앞서 비교한 서열들 중 차이가 있는 부위를 취합하여 각각 삼일열 원충과 열대열 원충의 LDH에 특이적인 epitope를 선정하였다.



▶▶ 그림 2. 삼일열 및 열대열 원충의 hydrophilicity prediction을 통한 epitope 예측.

A : 삼일열 원충 LDH, B : 열대열 원충 LDH

그 결과 아미노산 서열 110~120번에서 열대열 원충의 LDH는 hydrophilicity가 높아 epitope로 선정하였으나, 삼일열 원충의 LDH는 hydrophilicity가 낮아 epitope로 선정하지 않았다. 반면 205~220번에서는 삼일열 원충의 LDH는 hydrophilicity가 높아 epitope로 선정하였으나, 열대열 원충의 LDH는 hydrophilicity가 낮아 선정하지 않았다. 즉, 열대열 원충에서는 LDH의 아미노산 서열 110~120번(GGHIKKNCPNA)을, 삼일열 원충의 LDH는 205~220번 아미노산 서열(NNKKITDEEVEGIFDR)을 각각에 특이적인 epitope로 선정하였다.

III. 결론

본 연구에서는 삼일열 원충과 열대열 원충에서 각각 특이적인 단백질인 LDH의 서열상의 차이를 비교하였다. 이를 바탕으로 epitope 예측 프로그램을 이용하여 hydrophilicity를 분석하고 진단키트 제작에 사용할 수 있는 epitope를 예측하였다. 그 결과 아미노산 서열상에서 차이가 나는 부분과 예측된 epitope부분이 일치하는 것을 볼 수 있었다. 따라서 우리는 열대열 원충의 LDH에서는 110 ~ 120번의 아미노산 서열(GGHIKKNCPNA)을, 삼일열 원충의 LDH에서는 205 ~ 220번의 아미노산 서열(NNKKITDEEVEGIFDR)을 각각 말라리아 열원충에 특이적인 epitope로 선정하였다. 또한 이 두 서열을 말라리아의 다른 단백질들의 서열과 비교, 분석하였으나 유사한 서열이 없는 것으로 나타났기 때문에 epitope로써 성공 가능성이 매우 높다고 할 수 있다. 따라서 이 서열을 이용하여 삼일열 원충과 열대열 원충에 특이적인 말라리아 진단키트 개발에 사용한다면 좀 더 효율적인 진단이 가능할 것으로 사료된다.

■ 참고 문헌 ■

- [1] World Health Organization: media centre, Fact sheets, Malaria, 2014
- [2] J. M. R. Parker, "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*, 25, pp.5425-5432, 1986
- [3] Young B. Kim, "Immunogenicity and Ability of Variable Loop-Deleted Human Immunodeficiency Virus Type 1 Envelope Glycoproteins to Elicit Neutralizing Antibodies", *Virology*, 305, pp.124-137, 2003