pfhrp2/pfhrp3 유전자 결여 열대열 말라리아 특이 진단을 위한 생물정보학 기반 차세대 항원 단백질 선정

Selection of next-generation antigen protein for diagnosis of *pfhrp2/pfhrp3* gene deleted *plasmodium falciparum* based on bioinformatics

서 승 환, 이 지 후, 최 재 원, 김 학 용
충북대학교 생화학과

Seung Hwan Seo, Jihoo Lee, Jae-Won Choi, Hak Yong Kim
Chungbuk National University

요약

열대열 말라리아(*Plasmodium falciparum, P. falciparum, P. f*) 신속진단키트의 경우, *P. falciparum*에 특이적인 단백질로써 *Histidine Rich Protein 2 (PfHRP2)*가 사용되고 있다. 그러나 최근 연구에서 남아메리카와 중앙아메리카를 중심으로 *pfhrp2/pfhrp3* 유전자가 결여된 *P. falciparum* 열원충이 나타나는 것으로 보고된 바 있다. 본 연구에서는 생물정보학을 기반으로 *PfHRP2* 항원 단백질을 대체할 수 있는 새로운 *P. falciparum* 특이 항원 단백질을 선정하고자, PlasmoDB에서 5,777개의 *P. falciparum* 관련 단백질 리스트를 얻었다. 이후 NCBI BLAST를 통해 단백질 아미노산 서열을 분석하고 정상인에게 존재하지 않으며, 동시에 다른 말라리아 열원충(*P. vivax, P. ovale, P. malariae, P. knowlesi*)에도 존재하지 않는 *P. falciparum* 특이 아미노산 서열을 가진 단백질 15개를 추출하였다. IEDB analysis를 이용하여 에피토프, 수용성, 베타-턴, 접근성, 유연성, 면역원성을 분석하여 높은 평균값을 갖는 상위 3개 단백질을 선별하였다. KEGG pathway와 EMBL-EBI를 통해 선별된 3개 단백질의 혈액내 검출 가능성 및 아미노산 서열의 보존성을 분석하여 최종적으로 Glutamate-Rich Protein (GLURP)을 선정하였다. AIDA를 통해 단백질 아미노산 서열을 이용한 3차 구조 예측으로 GLURP의 구조 및 항체와의 결합을 도식화하였다. 최종적으로 선정한 GLURP는 *pfhrp2/pfhrp3* 유전자 결여 *P. falciparum*까지 특이적으로 진단이 가능하여 차세대 *P. falciparum* 특이 신속진단키트 개발에 도움이 될 수 있을 것으로 기대한다.

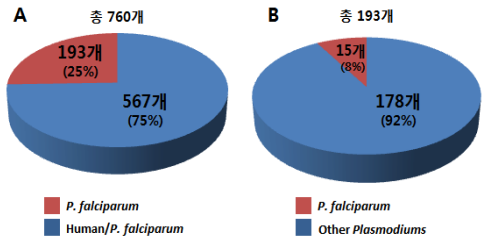
Ⅰ. 서론

전 세계에 걸쳐 PfHRP2를 이용한 열대열 말라리아(*Plasmodium falciparum*, *P. falciparum*) 특이 신속진단키트가 제작 및 판매되고 있다[1]. 그러나 최근 여러 연구에서 *pfhrp2/pfhrp3* 유전자가 결여 된 *P. falciparum* 열원충이 나타나는 것으로 보고된 바 있어, 더 이상 PfHRP2 기반 신속진단키트만으로는 *P. falciparum* 특이 진단이 불가능하다[2]. 따라서 우리는 PlasmoDB에서 *P. falciparum* 관련 단백질 리스트를 얻고 BLAST를 통한 단백질 아미노산 서열 비교분석을 통해 *P. falciparum* 특이 단백질들을 선별하였다. 더 나아가 KEGG pathway와 EMBl-EBI를 활용하여 *P. falciparum* 특이 단백질들의 특성을 분석하였다. 최종적으로 현재 열대열 말라리아 진단에 활용되고 있는 PfHRP2를 대체하여, 보다 정확하게 *P. falciparum* 특이 신속진단키트 제작을 위한 차세대 항원 단백질로 GLURP를 제시하고자 한다.

Ⅱ. 본론

1. *P. falciparum* 열원충 관련 단백질 추출 및 *P. falciparum* 특이 단백질 선별

PlasmoDB를 통해 *P. falciparum* 열원충 관련 단백질 총 5,777개 중, 불확실한 정보를 가진 단백질들과 PfHRP2, PfHRP3를 제외한 760개의 단백질 리스트를 얻었다. 이후 BLAST를 이용하여 각 단백질들의 아미노산 서열 비교 분석을 통해 정상인 그리고 *P. falciparum* 열원충을 제외한 다른 4종류의 말라리아 열원충에 존재하지 않는, *P. falciparum* 특이 단백질 후보 15개를 선별하였다[그림 1].



▶▶ 그림 1. BLAST를 이용한 *P. falciparum* 특이 단백질 후보군 선별

(A) *P. falciparum* 관련 단백질 760개 중, 정상인(567개)에는 존재하지 않는 단백질 193개 선별. (B) 193개 단백질 중 사람에게 감염을 유발하는 다른 4종류 열원충에 존재하지 않는 *P. falciparum* 특이 단백질 후보 15개 선별. Other *Plasmodiums*; *P. vivax, P. ovale, P. malariae, P. knowlesi.*

2. IEDB analysis를 통한 상위 3개 *P. falciparum* 특이 단백질 후보군 선별

IEDB analysis는 단백질 아미노산 서열을 이용하여 에피토프(epitope), 수용성(hydrophilicity), 베타-턴(beta-turn), 접근성(accessibility), 유연성(flexibility), 면역원성(immunogenicity)을 예측하는 프로그램이다. 앞서 선별한 15개의 후보 단백질 중, 각 method 별 평균값을 모두 더하여 높은 값을 갖는 단백질 순으로 분류하고, 8점 이상의 상위 3개 단백질 MSP2, STARP, GLURP를 선별하였다[표1].

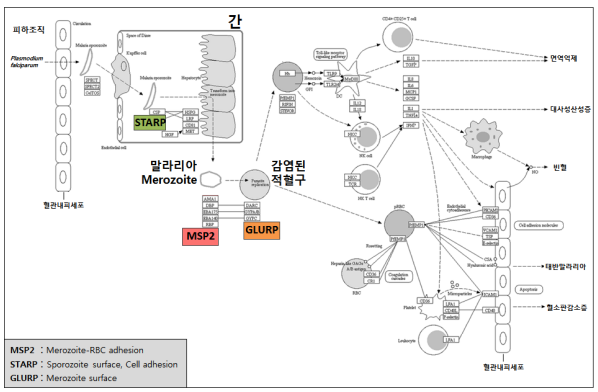
표 1. IEDB analysis를 통한 15개 *P. falciparum* 특이 단백질 후보중 상위 3개 단백질 선별

Name	Epitope	Hydrophilicity	Beta-turn	Assecibility	Flexibility	Antigenicity	Total
MSP2	1.035	3.779	1.101	1	1.046	0.974	8.935
STARP	0.655	3.568	1.112	1	1.041	0.950	8.326
GLURP	0.782	3.235	0.989	1	1.033	0.996	8.035
GBP	0.692	3.216	1.007	1	1.026	0.981	7.922
MSP11	0.703	3.144	1.063	1	1.034	0.970	7.914
PF332	0.596	3.190	0.910	1	1.027	1.000	7.723
H101	0.520	2.867	1.042	1	1.029	0.986	7.444
EBA175	0.380	2.727	1.035	1	1.022	0.992	7.156
EBA140	0.328	2.633	1.053	1	1.020	0.992	7.026
MSRP7	0.285	2.475	1.039	1	1.023	0.997	6.819
EPF3	-0.052	1.580	1.009	1	1.007	1.000	5.544
RIF	-0.184	1.371	0.940	1	0.993	1.039	5.159
STEVOR	-0.301	1.130	0.949	1	0.990	1.037	4.805
MAHRP2	-0.270	0.974	0.933	1	0.973	1.068	4.678
MC-2TM	-0.572	0.478	0.948	1	0.988	1.027	3.869

Score=각 단백질의 method 별 평균값. 각 method 별 평균값의 총 합을 이용, 가장 높은 값을 가진 순부터 정렬

3. KEGG pathway를 이용한 상위 3개 단백질의 위치 및 기능 분석

KEGG pathway를 통해 *P. falciparum* 감염경로에 따른 상위 3개 단백질의 위치 및 기능을 확인하였다[그림 2].



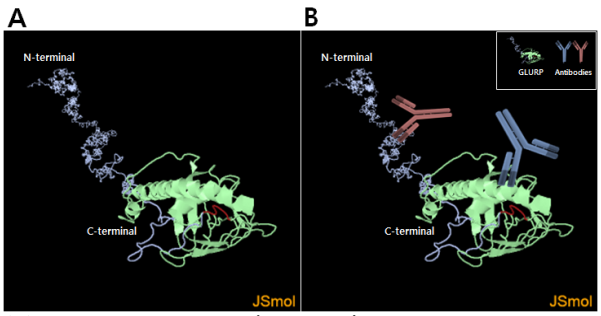
▶▶ 그림 2. KEGG pathway를 통한 단백질 위치 및 기능
주황색=GLURP, 분홍색=MSP2, 녹색=STARP

4. 상위 3개 단백질 특성 분석 및 최종 단백질 선정

앞서 선별 된 단백질을 이용한 신속진단키트 제작에 있어서 무엇보다 중요한 부분은 진단 시 항원 단백질의 검출 가능성이다. STARP의 경우, KEGG pathway 분석 결과 *P. falciparum* 열원충의 감염 후 잠복기 상태인 sporozoite에 위치함으로 혈액을 이용한 진단 시 검출 가능성이 낮다[그림 2]. 또한 EMBL-EBI를 통해 조사해본 결과, MSP2의 경우에는 아미노산 서열의 보존성이 낮은 variable region이 존재하기 때문에 계속적인 구조적 변화를 초래할 가능성이 높다. 이를 바탕으로 STARP와 MSP2를 제외한 Glutamate-Rich Protein (GLURP)을 *P. falciparum* 특이 단백질로 최종 선정하였다.

5. GLURP 3차 구조 예측 및 항체와의 결합 모식도

최종 선정한 GLURP의 구조적인 형태와 항-GLURP 항체와의 결합 가능성을 분석하기 위해 단백질 아미노산 서열을 이용, Ab Initio Domain Assembely(AIDA)를 통해 GLURP의 3차 구조를 예측하였다[그림 3A]. 더 나아가 신속진단키트 적용을 위해 GLURP와 항-GLURP 항체와의 결합을 예측해본 결과, 안정적인 항원-항체 결합을 이룰 것으로 예측 되었다[그림 3B].



▶▶ 그림 3. GLURP의 3차 구조 예측 및 GLURP modeling
(A) 예측된 GLURP 단백질 3차 구조. 파란색=N-말단, 녹색=C-말단. (B) GLURP와 항체 결합 예상 모식도.

Ⅲ. 결론

최근 *pfhrp2/pfhrp3* 유전자 결여 *P. falciparum* 열원충이 등장함에 따라 기존의 신속진단키트에 이용되고 있는 PfHRP2를 대체할 수 있는 단백질의 필요성이 점차 대두되고 있다. 본 연구에서는 생물정보학을 기반으로 BLAST를 통한 단백질 아미노산 서열을 비교 분석하고, IEDB analysis를 이용, 각 method 별 평균값의 총 합을 계산하여 *P. falciparum* 특이 단백질 상위 3개를 선별하였다. KEGG pathway와 EMBL-EBI를 통해 선별한 단백질들의 신속진단키트 적용 시, 혈액 내 검출 가능성과 아미노산 서열의 보존성을 고려하여 항체와의 안정적인 결합 가능성을 분석한 후 GLURP를 *P. falciparum* 특이 단백질로 최종 선정하였다. 본 연구결과를 통해 제시한 GLURP는 PfHRP2를 대체 할 차세대 항원 단백질로, *P. falciparum* 특이 신속진단키트 개발에 도움이 될 수 있을 것이다.

▌참 고 문 헌▐

[1] Nelson Lee, "Identification of Optimal Epitopes for Plasmodium falciparum Rapid Diagnostic Tests That Target Histidine-Rich Proteins 2 and 3", JCM, VoL 50, No. 4, pp. 1397-1405, 2012

[2] Sheila A. Okoth, "Variation in Plasmodium falciparum Histidine-Rich Protein 2 (Pfhrp2) and Plasmodium falciparum Histidine-Rich Protein 3 (Pfhrp3) Gene Deletions in Guyana and Suriname", PLoS One, VoL 10, No. 5, e0126805, 2015