

# easyLMO : LMO 실험을 위한 유전자 편집 Editor 설계 및 구현

## easyLMO : Implementation of Gene Editor for LMO

최한석, 김종화, 김동욱\*  
 목포대학교 컴퓨터공학과, (주)올비티\*

Han Suk Choi, Jong Hwa Kim, Dong-Wook Kim\*  
 Department of Computer Engineering Mokpo National  
 University, AIIDTec co. Ltd\*

### 요약

본 논문에서 구현한 시스템은 LMO를 생산하기 위하여 수행하는 실험과정 중 Protein coding을 변화시켜 인위적으로 변형된 생물체를 확보하기 위하여 수행하는 전처리 작업을 용이하게 수행 할 수 있도록 설계되고 구현되었다.

## I. 서론

LMO(Living Modified Organisms)는 현대생명공학기술을 이용하여 생물종의 유전물질을 인위적으로 변형시킨 생물체를 포괄적으로 지칭하며, 농업분야에서는 농산물의 생산량 증대, 품질 향상 등을 목적으로 주로 사용되고 있으며 유전자변형기술을 통해 개발된 작물은 1996년 상업화 이후 지속적으로 증가하고 있다[1].

이는 1996년 유전자변형작물의 재배면적 170만 ha에 비해 약 100배 이상의 증가세를 보임으로써, 근래 농업 역사에 있어 가장 빠르게 채택된 작물 생산 기술로 평가된다. 그러나 LMO를 만들수 있는 과정 중 다양한 LMO 변형을 시도하여야 하는데 자동화 된 과정이 없어 연구원이 수많은 protein code를 눈으로 확인하고 이를 수정한 후 실험에 적용하고 있다[1].

이에 본 논문에서 제안하는 시스템은 밝혀진 protein을 자동으로 찾아내고 유사종에 적용가능한 LMO 조합을 자동으로 찾아내어 시퀀스를 변형하여 연구자에게 제공함으로써 유전자 편집 기술의 수월성을 높이도록 설계되고 구현되었다.

## II. 시스템 설계 및 구현

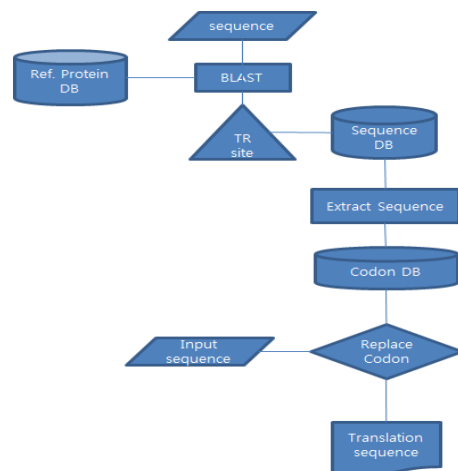
### 1. 시스템 고려사항

본 시스템의 설계시 고려사항은 미지의 종 특이적인 protein coding 빈도를 확보하고 이를 활용하여 새로운 protein 조합을 생성하며 정확한 protein Translation Frame을 확보하기 위하여 기 발견된 유사종의 Protein과 BLAST하여 Translation Site를 확보하여 오번역을 방지할 수 있도록 하여야 한다. 또한, 확보된 Translation site를 이용하여 입력된 sequence를 다양한 조건으로 파싱해야 하며 파싱된 시퀀스를 Translation 시킨 후 빈도 계산을 수행한 후 계산된 빈도를 데이터베이스에 저장, 로드, 삭제 할 수 있는 모듈이 개발되어야 한다. 이후 계

산된 빈도를 불러온 후 새로운 시퀀스를 계산된 빈도를 참조하여 자동 수정한다. 이때 수정된 빈도내에서 선택된 enzyme site를 회피 할 수 있어야 하며 Protein database를 다운로드 받아 자동으로 BLAST DB화 시킬 수 있어야 한다.

### 2. 시스템 설계

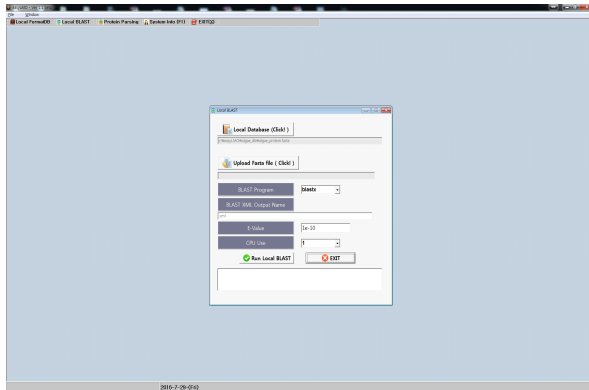
easyLMO는 레퍼런스 BLAST DB를 확보한 후 BLAST를 수행 한 후 시퀀스를 파싱 데이터베이스에 저장한다. 저장된 데이터를 이용하여 입력된 시퀀스에서 그림 1와 같은 조건을 체크하여 조건에 맞는 시퀀스를 확보한다. 확보된 시퀀스를 이용하여 데이터베이스에 분석된 정보를 저장하고 새로 입력된 시퀀스를 코돈테이블의 상위 코돈으로 치환하고 선택된 enzyme site를 회피할 수 있도록 구성되었다. 이러한 최종 과정을 거치게 되면 최종 sequence를 확보할 수 있도록 설계되었다. 그림 1은 easyLMO 시스템의 데이터 흐름도이다.



▶▶ 그림 1. easyLMO 데이터 흐름도

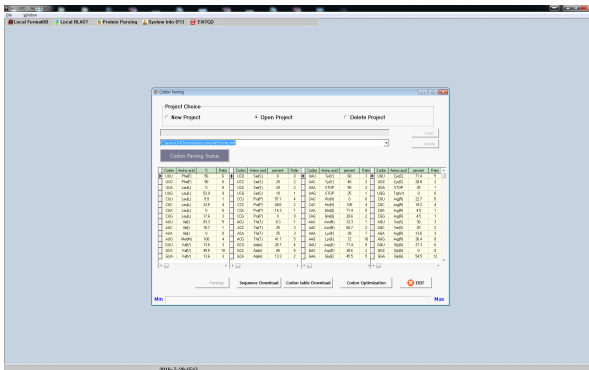
## 2. 시스템 구현

그림 2, 그림 3, 그림 4는 easyLMO 시스템 구현 화면이다. 그림 2는 확보된 레퍼런스 데이터베이스에 확보된 시퀀스를 입력하여 BLAST 검색을 하여 Translation site를 확보할 수 있는 과정이다.



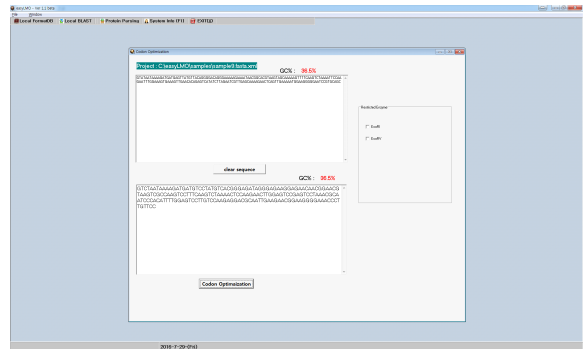
▶▶ 그림 2. Translation 정보를 위한 BLAST

그림 3은 BLAST된 결과파일을 이용하여 자동으로 시퀀스를 Translation 한 후 각각의 빈도 값을 계산하여 데이터베이스에 저장할 수 있으며 이전 과정을 통해 확보된 코돈테이블을 관리할 수 있도록 구현되었다.



▶▶ 그림 3. Codon Table 로딩

그림 4은 입력된 시퀀스를 코돈테이블을 참조하여 높은 빈도 및 enzyme site를 회피하여 원하는 시퀀스를 얻을 수 있도록 구현되었다.



▶▶ 그림 4. High value sequence 교체

## IV. 결론

본 연구에서는 밝혀진 protein을 자동으로 찾아낸 후, 유사종에 적용가능한 LMO 조합을 자동으로 찾아내어 사용자의 수고 없이 시퀀스를 변형하여 연구자에게 제공할 수 있는 easyLMO 시스템을 설계하고 구현 하였다. easyLMO는 유전체 연구자들에게 유전자 편집 기술의 수월성을 높이도록 구현되었다.

본 성과물은(논문, 산업재산권, 품종보호권 등)은 농촌진흥청 포스트게놈다부처유전체연구사업 (PJ013693032018)의 지원에 의해 이루어진 것임

## ■ 참고 문헌 ■

- [1] Jurka J, Kapitonov VV, Pavlicek A, Klonowski P, Kohany O, Walichiewicz J: Repbase Update, a database of eukaryotic repetitive elements. Cytogenet Genome Res 2005, 110:462-467.
- [2] Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: Basic local alignment search tool. J Mol Biol 1990, 215:403-410.