

RIA 기반 DNA서열 분석도구의 설계 및 구현

The Design and Implementation of RIA-Based DNA Sequence Analysis Tools

김명관*, 조충효**

Myung-Gwan Kim, Choong-Hyo Cho

요 약 생명정보학 분야의 발전에 따라 방대한 양의 DNA서열 데이터를 효율적으로 분석하기 위해 분석도구가 사용되고 있다. 하지만 기존의 분석도구들은 분석하고자 하는 데이터를 찾고, 적용해야 하는 불편함이 있다. 본 논문에서는 이러한 문제점을 해결하기 위하여 웹2.0기반 RIA(Rich Internet Application) 방식으로 구현한 분석도구를 제안한다. RIA방식을 적용한 분석도구는 기존 웹 방식의 문제점을 보완한 웹2.0기반에서 DNA서열 데이터를 찾고, 실시간으로 분석내용을 보여준다. 개발된 웹 애플리케이션은 윈도우 시스템 상에서 Flex2를 이용하였다.

Abstract Due to the progress of Bioinformatics field, We are making use of analyzing tools for effective analyzing enormous data of DNA sequence. But there was inconvenience in existing tools when searching and applying data for analyzing. Our treatise proposes a tool developed by a form based on RIA(Rich Internet Application) that you can solve the problems came from weak points. The analyzing tool for RIA indexing data of DNA sequence shows the results by real time in basis of Web 2.0 which supplemented basis on a form of Web. The web application was developed in Flex2 on Windows workstation.

Key Word : Bioinformatics(생명정보학), Web2.0(웹2.0), RIA(Rich Internet Application)

I. 서론

유전체 서열과 기능을 알아내기 위한 인간유전체산업(Human Genome Project, HGP)이 국제협력 사업으로 1990년에 시작되어 1999년에 완료 되었다. 현재 바이러스, 식물, 동물 등 지구상에 존재하는 생물체의 유전체 사업이 전 세계적으로 진행되고 있으며, 1995년 Haemophilus influenzae, RD[1] 의 전체 유전체 서열결정을 시작으로, 현재 생물학의 모든 분야에 대한 서열정보가 대량으로 늘어나고 있다.[2]

이에 따라 연구자가 다루는 데이터는 비약적으로 증

가되어 데이터를 수작업으로 처리한다는 것은 매우 어려운 일이 되었다. 이런 상황에서 바이오 분야의 연구개발의 효율을 높이기 위한 정보기술로서 생명정보학(Bioinformatics)이 더욱더 중요해지고 있다.[3]

실제로 생물 정보학의 여러 분야 중에서 서열을 분석하고, 유전자 기능을 예측하는 여러 가지 분석도구가 개발되어 유전체 연구에서 핵심적인 역할을 수행하고 있다.[4] 유전자 서열 분석관련 도구들의 종류로는 Artemis와 Phred, Phrap 프로그램 등이 있다. Phred와 Phrap은 염기서열 분석의 가장 대표적인 소프트웨어이고, Artemis는 영국의 생어 연구소(Sanger Institute)에서 자바로 개발한 서열 분석도구 이다.[5][6][7] 그러나 기존 유전자 서열 분석관련 도구들은 리눅스나 유닉스 기반의

*정회원, 을지대학교 의료산업학부

**준회원 을지대학교 컴퓨터정보과

접수일자 2009.03.10, 수정완료일자 2009.04.03

프로그램이기 때문에 결과물을 얻어내기 위해서 프로그램을 서버에 설치하고 명령어 사용법과 각종 옵션을 숙지해야만 한다는 단점을 가지고 있다. 최근 웹으로 제공되는 분석도구가 있지만 기존 웹 방식은 단방향의 성격이 강하다. 사용자는 이미지와 텍스트로 이루어진 웹 페이지를 하이퍼링크를 통해 이동하면서 정보를 제공받기 때문에 대량의 서열을 분석하기 위해서는 많은 시간이 걸린다. 따라서 대량의 데이터를 빨리 처리할 수 있고, 사용자가 편리하게 작업을 수행할 수 있는 도구의 개발이 필요한 실정이다.[8][9]

본 논문에서는 유전자 분석에 필요한 분석도구를 웹 2.0기반 RIA(Rich Internet Application) 개발방식으로 구현한다. 웹으로 이루어진 프로그램은 사용자에게 브라우저만 있으면 이용 가능하여 프로그램을 배포하는 비용과 시간을 줄일 수 있다. 또한 기존의 HTML의 단점을 보완한 RIA(Rich Internet Application)방식을 사용하면 필요한 데이터만 호출하여 페이지 로딩 시간을 줄이고, 대량의 데이터를 빠르게 처리할 수 있다.[10][11]

논문의 구성은 제 2장에서 기존 웹 환경과 RIA(Rich Internet Application)의 특징 및 장점 대해 설명한다. 제 3장에서는 웹 2.0기반 RIA(Rich Internet Application) 개발방식으로 구현한 DNA서열 분석도구의 구현에 대해 설명하고, 제 4장에서 결론 및 향후 과제를 제시한다.

II. 관련연구

여기서는 기존의 DNA서열 분석도구들을 소개하고, 문제점에 대해 논의 한다. 그리고 기존의 문제점을 보완하기 위한 웹 2.0의 핵심 기술인 RIA(Rich Internet Application)에 대해 살펴본다.

1. 기존의 DNA서열 분석도구 사례

염기서열 분석의 가장 대표적인 소프트웨어에는 Phred와 Phrap가 있다. Phred는 DNA 서열로부터 생성된 데이터를 읽어 들여, 각각의 해당 염기를 결정해 주는 프로그램이다. [그림1] Phred는 SCF(AVI model 373 및 377), ESD(MegaBACE) 등의 크로마토그램 파일을 자동으로 읽어 들여 각각의 서열 정보를 FASTA, XBAP, PHD 형식등의 파일로 출력하며, 각 염기에 대한 품질 값을 나타내 주고 이 결과를 출력해 준다. Phred는 염기호출의 높은 정확성과 사용상의 편리, 각 염기에 대한 자세

한 품질 값 출력, 대학 사용자를 위한 자유 라이선스 등의 장점 때문에 현재 세계적으로도 가장 널리 이용되고 있다.

Phrap은 Phred와 함께 하나의 패키지처럼 구성되어 널리 알려진 가장 대표적인 서열 어셈블러(배열기;sequence assembler)로서, 인간 유전체 프로젝트에서도 Phrap은 그 진가를 발휘 하였다. Phrap의 서열 배열(sequence assembly) 속도는 매우 빠르고, 일반적인 유닉스 워크스테이션이나 PC에서 실행이 가능하다. [12]

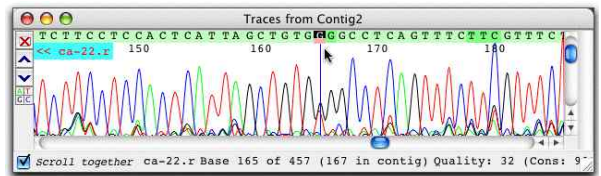


그림 1. Phred 프로그램 서열분석 화면
Fig 1. Screen Shot of Phred

추가적으로 인터넷을 이용하여 사용할 수 있는 Artemis 소프트웨어가 있다. Artemis는 Sanger Center에서 개발된 도구로써 서열분석기능과 함께 EMBL와 GenBank, GFF형식을 읽어서 그 내용을 그래픽으로 표현해 준다. <그림 2> 와 같이 6개 프레임으로 구성되어 있고 각각의 프레임은 서열을 분석결과와 주석내용을 작성하고 결과를 보여 준다. 또한 서열을 구성하고 있는 염기(A/C/G/T)서열을 그래프로 표현 하여 각각의 비율을 한눈에 분석 할 수 있게 구현 하였다. 서열을 막대로 표현하고 색으로 구분하여 색만으로도 서열의 특징이 무엇인지 구분 지어 놓았다.[13]

Artemis는 막대로 표현한 서열들을 염기서열로 표현하는 프레임을 따로 만들어 염기서열을 분석하는 위치를 알 수 있도록 구현 되어 있다. 마지막 프레임은 막대 색의 정의와 서열의 위치, 주석을 표현한다.[14]

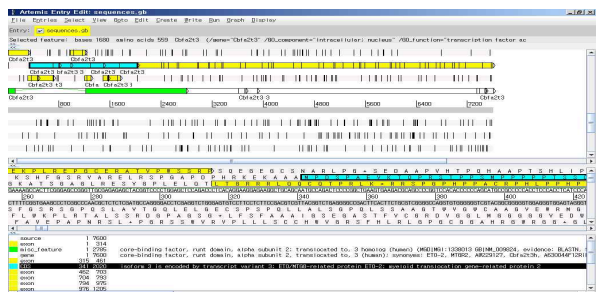


그림 2. 6개의 프레임으로 구성된 Artemis의 전체 화면
Fig 2. Screen Shot of Artemis

위와 같이 기존의 소프트웨어 분석도구들은 대부분 리눅스나 유닉스 기반 프로그램이어서 이를 설치하고 사용하는 데 어려움과 불편함이 많은 실정이다. 근래에 웹 기반 분석도구들이 개발되고 있으나 HTML로 만든 웹서비스는 페이지 단위 기반 구성으로 되어 있고, 계층적으로 연결되어 있어 필요 없는 데이터 받아야 하는 문제점이 있다.

2. RIA(Rich Internet Application)의 특징 및 장점

HTML(Hyper Text Markup Language) 기반에서의 하이퍼링크를 통한 페이지들의 연결의 편리성은 WWW(World Wide Web)을 짧은 시간에 전 세계적으로 보급시키는데 결정적인 역할을 하였다. 그러나 이러한 페이지 단위의 연결과 실행은 동기식으로 데이터를 처리함으로써 원하는 데이터까지 서버 측에 요구하여 서버 측 자원을 소비 하였다. 즉, 페이지가 로드될 때마다 내가 원하는 데이터뿐만 아니라 원치 않는 데이터들까지도 서버로부터 받게 되고, 사용자들은 데이터가 로드될 때까지 로딩 중인 하얀 화면의 웹 브라우저를 바라보고 있어야 했다. 하지만 RIA(Rich Internet Application) 개발 방식은 페이지의 재 호출 없이 내가 원하는 데이터를 서버에 요구하고 바로 받아서 처리 할 수 있다. 현재 이용할 수 있는 대한 극장 온라인 예매 시스템(<http://www.daehancinema.co.kr/>)의 경우, 기존의 예약 처리 프로세스는 약 5~6단계를 거쳐 이루어 졌지만, RIA(Rich Internet Application)방식으로 개발한 후에는 한 화면에서 한 번에 처리할 수 있게 되었다. 결과적으로 예매 처리 시간이 기존보다 10배 이상(5분→30초) 빨라졌고, 그로 인해 고객 예매율은 증가하였다.

본 연구에서는 사용자의 편의성과 분석 효율성을 높이기 위해 유전자 분석과정에서 여러 단계를 거쳐야 했던 기존의 웹의 문제점을 해결할 수 있는 웹2.0 기반의 RIA(Rich Internet Application)개발방식을 제공하고자 한다. RIA(Rich Internet Application)의 클라이언트환경은 프로세싱 및 데이터 분리를 통해 기존 웹 환경의 문제점을 해결 할 수 있다.

RIA(Rich Internet Application)개발 방식은 페이지의 재 호출 없이 원하는 데이터를 서버에 요구하고 바로 받아서 처리할 수 있다. 또한 실시간으로 서버와 데이터 교환이 가능하고, 다양한 사용자 컨트롤과 그래픽 효과를

개발할 수 있도록 컴포넌트 라이브러리를 제공한다. 이러한 RIA(Rich Internet Application)의 장점을 가지고 새로운 분석도구를 구현하면 유전자 분석 효율성을 증진시킬 수 있다.<표 1> RIA(Rich Internet Application)개발 방식에는 어도비의 Flex, Ajax, WPF, Lazzlo, Droplet 등이 있다. 특히, Flex는 개발 생산성, 애플리케이션 활용, 애플리케이션 성능 향상의 효과를 극대화한 RIA(Rich Internet Application)개발 방식이다. 본 연구에서 RIA(Rich Internet Application)개발 방식에서 가장 널리 알려진 어도비의 Flex2로 DNA서열 분석도구를 구현 하였다.[13]

표 1. RIA의 특징 및 장점
Table 1. Features of RIA

HTML의 특징 및 한계	기존의 Flash 구성의 한계	Rich Internet Applications
· 페이지 전체의 구성요소를 로딩해야 함	· 파일을 한번에 로딩하기 때문에 시간이 많이 걸림	플래시 인터페이스를 사용해 전체페이지가 아닌 필요한 부분만 로딩해 페이지 부하가 적고 빠름
· 프로세스에 따라 연결된 페이지들을 모두 로딩 해야 함	· 검색이나 DB연동이 되지 않는 단순 멀티미디어로 사용됨	비주얼적 다이나믹 검색과 DB연동이 가능하고 콘텐츠를 재가공 없이 모바일에서 사용가능
· 멀티미디어 콘텐츠의 수용에 제한이 있음	· HTML에 비해 제작기법이 복잡해 유지보수가 까다로움	복잡한 단계별 페이지 구조를 단일 Interface에서 처리 가능

III. 웹2.0기반 DNA서열 분석도구의 구현

1. 개발목적

기존 서열분석 프로그램을 사용하기 위해서는 필요한 서열분석도구를 찾고, 설치를 하는 불편함 있다. 이러한 불편함을 해결하고자 웹 환경에서 서열분석도구를 제공한다면 사용자들이 인터넷 접속만으로 서열분석도구를 사용할 수 있다.

또한 기존에 제공되고 있는 GenBank 파일을 보면 많은 정보들이 들어있지만, Text 형태로 데이터를 보여주고 있어 한눈에 알아보기 어렵다. 그리고 일반 사람들이 보기에는 무슨 내용을 포함하고 있는지 알기 힘든 문제

점을 발견하였다. 이러한 문제점을 보완 하기위하여 Text 기반의 서열분석을 RIA(Rich Internet Application)의 강점인 풍부한 콘텐츠 사용하여 정렬된 서열분석을 보여준다면 서열 분석의 효율성을 높일 수 있다.

2. DNA서열 분석도구의 화면구성 원리

HTML 방식보다 역동적인 화면 연출이 가능하며 화면이 바뀔 때마다 페이지를 새로 고쳐야 했던 문제점을 RIA(Rich Internet Application)를 통해 한 페이지 내에서 모든 정보 이용이 가능하기 때문에 웹 사용성도 비약적으로 높일 수 있는 장점이 있다. <그림 4>과 <그림 5>는 HTML으로 구현했을 때와 RIA(Rich Internet Application)로 구현한 화면을 비교한 것이다. <그림 4>는 미국 국립생물정보센터(<http://www.ncbi.nlm.nih.gov>)에서 유전체 정보 데이터베이스인 GenBank파일을 검색하기 위해 4차례 페이지 이동을 하였지만, <그림 5>는 한 페이지에서 모든 화면을 보여줄 수 있다. 그 결과 여러 GenBank파일을 검색할 경우 HTML로 만든 사이트에 비해 많은 시간과 노력을 절감 할 수 있다.

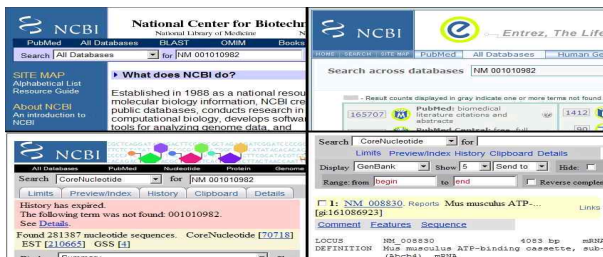


그림 4. HTML로 구현한 NCBI사이트 에서 GenBank파일 검색
Fig 4. Search for GenBank File from NCBI Site



그림 5. RIA로 구현한 GenBank파일 검색 및 분석
Fig 5. Search and Analysis for GenBank File by RIA

3. GenBank파일의 DNA서열 분석과정

생물체가 살아있기 위해서는 반드시 정해진 유전 정보의 발현에 달려있다. 중심원리(Central Dogma)로 불리는 이 명제는 전체 생명정보의 흐름을 가장 단순화 시켜 이해할 수 있도록 한다. 유전정보의 흐름은 DNA에서 RNA로 RNA에서 번역(translation)을 통해 단백질서열로 만들어진다. 번역이란 유전자 부분의 DNA서열이 RNA 서열로 전사(transcription)된 후 이를 바탕으로 단백질 서열을 만드는 과정이다. 번역과정에서 RNA 서열 자체가 단백질 서열로 번역 되지는 않는다. 세포내에서 단백질 서열을 만드는 기능을 가진 효소들이 RNA 서열, 특히 전령 RNA(mRNA) 서열을 읽으면서 이미 세포내에 존재하는 개개의 아미노산 분자들을 순서에 맞게 연결하여 단백질 서열을 만드는 것이다.

본 연구에서는 GenBank의 DNA서열을 가지고 번역한 단백질 서열을 <그림 6> 과 같이 표현하였다. DNA서열에 번역과정을 통한 단백질서열의 표현으로 서열분석을 한 눈에 빠르게 분석할 수 있다.



그림 6. DNA서열번역과 아미노산의 표현
Fig 6. Representation of DNA and Amino Sequence

4. 웹2.0기반 DNA서열 분석도구의 구현

본 연구에서 구현한 DNA서열 분석도구는 <그림 7>에서 보는 바와 같이 DNA서열검색기능, 기본서열정보, 제한효소분석, 서열분석기능으로 구성된다. 분석도구를 설치할 필요 없이 웹에 접속하여 필요한 DNA서열 데이터를 찾으면 바로 서열분석 결과를 보여준다.

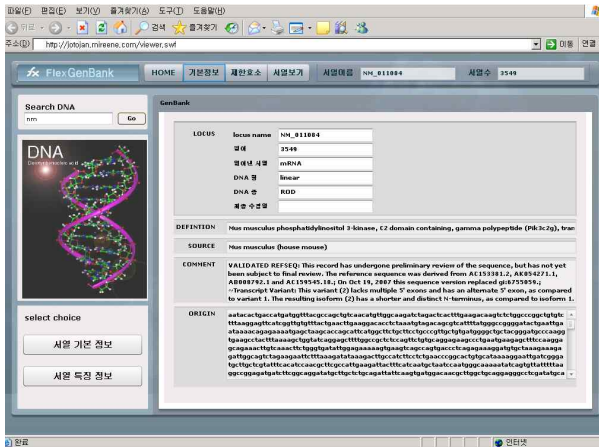


그림 7. Web2.0기반 서열분석 도구 전체화면
Fig 7. Screen Shot of Our System

4.1 GenBank에서 DNA서열 검색

일반적으로 서열분석 도구는 분석하고자 하는 서열과 일을 여러 데이터 제공기관들 (NCBI, UCSC, EBI 등)에서 서열 검색 서비스를 통해 다운받아 <그림 8>와 같이 분석도구에 적용하는 방식이다. 여러 데이터 제공기관들은 기존의 웹 방식으로 서비스를 하고 있어 서열을 검색할 때 마다 새로운 페이지로 이동한다. 이러한 방식은 대용량의 서열파일 검색에 많은 시간이 필요하게 된다. 이러한 문제점을 해결하고자 <그림 9>와 같이 분석도구 안에 검색창을 구현함으로써 서열검색과 동시에 분석결과를 보여 준다. 이러한 결과는 서열파일의 검색 시간을 감소시켜 사용자에게 편리함을 제공한다.



그림 8. Artemis 서열분석 도구에 서열파일 적용
Fig 8. Open Menu of Artemis

Web2.0기반 DNA서열 분석도구는 NCBI에서 제공되는 GenBank의 XML파일을 웹서비스로 함수를 호출하고 그 결과를 SOAP(Simple Object Access Protocol)메시지로 받는 방식이다.



그림 9. 분석도구의 서열검색 기능 및 검색결과
Fig 9. Search for DNA Sequence

4.2 GenBank파일의 기본서열 분석

서열 분석 과정에서 서열의 특성 및 서열이 가진 유전자의 위치 및 기능 등을 분석하여 주석을 작성 하여 다른 연구자들도 알 수 있도록 약속된 형식의 파일로 정리하는데, 그 대표적인 형식이 GenBank 파일 <그림 10> 이다.[12] 서열 기본정보에서 기존에 TEXT형식의 GenBank파일을 파싱하여 분류하였다. 이와 같은 분류는 사용자에게 서열분석의 내용을 쉽게 이해할 수 있게 표현 한 것이다.

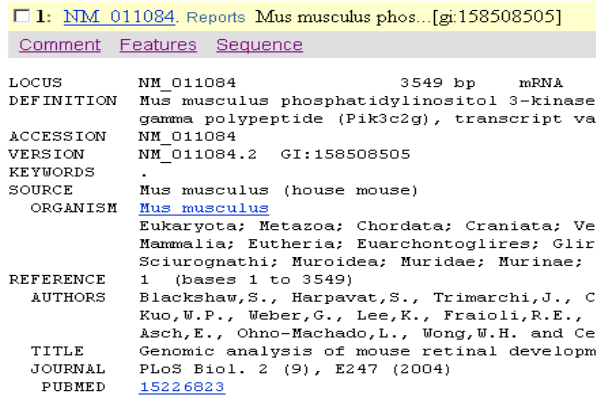


그림 10. NCBI의 GenBank 파일
Fig 10. GenBank File of NCBI

Web2.0기반 DNA서열 분석도구에서 기본정보는 DNA서열을 알려주고, 다른 연구자의 서열에 대한 연구 내용을 COMMENT 부분을 통해 <그림 11> 알 수 있다.



그림 11. RIA환경에서의 GenBank파일을 파싱한 화면
Fig 11. GenBank Parsing on RIA Tools

4.3 제한지도

서열분석 작업 중에 제한지도(restriction map)작성이 있다. 제한지도는 전체 염기서열에서 제한효소가 인식하는 서열의 위치를 찾아서 정리한 것이다. 제한효소(restriction enzyme)는 전체 염기서열에서 특정 서열(인식 사이트, recognition site)을 찾아서 특정 서열내의 특정 위치를 절단하는 기능을 가진 효소이다. 목록에서 제한효소를 선택하고 적용하면 [그림 12] 에서와 같이 적용한 제한효소에 의해 절단된 위치를 나타낸다.



그림 12. 서열에 제한효소 EcoRI를 적용한 제한지도
Fig 12. Restriction Map

4.4 서열분석 기능

모든 생명현상은 DNA서열에 들어 있는 정보에 의해 일어나고, 그러한 정보는 서열을 여러 가지 방법으로 조작해야 얻을 수 있다. 서열의 조작 과정에서 효율적인 분류를 위해 <그림 13>에서 보는 바와 같이 서열을 한 줄로 배치하였다. 이러한 방식은 서열 분석을 일괄적으로 확인할 수 있으므로 서열분석의 생산성을 높일 수 있게 된다.[15]

왼쪽화면에는 서열의 특징과 위치를 따로 구현하였다. 서열의 어느 부분이 어떤 유전자의 서열인지, 그 유전자의 이름 및 기능은 무엇인지 등 특정 영역에 대한 실질적인 분석결과가 표시된다. 서열 데이터에서 특정 영역에 대한 정보(Feature)를 표시하는 것을 주석이라고 하며 생체 내에서 DNA가 매우 많은 역할을 하듯이 이 내용을 표시하기 위한 feature-type들도 source, gene, CDS 등 많은 종류가 있다. 이러한 정보의 위치를 효율적으로 구분 할 수 있게 스크롤바를 이용하였다. 서열분석의 특성에는 서열을 구성하고 있는 각 염기(A/C/G/T)의 비율이 있다. GC-contents는 계놈의 전체 염기서열이나 특정 영역의 부분 서열에서 G와 C가 차지하는 비율을 말한다. RIA(Rich Internet Application)환경에서 서열분석 도구는 <그림 14>과 같이 그래프를 이용하여 전체 GC-contents를 한번에 파악 하고, Text의 표현으로 정확한 수치를 알 수 있다.[16]

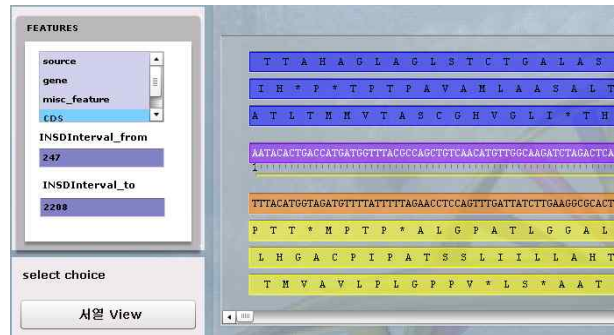


그림 13. 일괄적인 서열분석을 위한 인터페이스
Fig 13. UI for DNA Analysis



그림 14. GC-contents 인터페이스
Fig 14. GC-Contents Interface

IV. 결론 및 향후 연구

본 연구에서는 유전체 사업의 수행에 필수적인 소프트웨어를 웹2.0 기반 RIA(Rich Internet Application)로 구현하여 생물학자들의 편리성과 분석의 생산성을 높인 분석도구를 제시하였다.[20]

지금까지 서열 분석도구는 프로그램을 설치하고, 프로그램에서 요구하는 입력 형태로 데이터를 제공해야 하므로 컴퓨터에 대해 전문적인 지식이 부족한 생물학자 입장에서는 어려운 일이 아닐 수 없다. 또한 기존 웹 환경에서 제공된 분석도구는 페이지 로딩 시간이 많아 대용량 서열 데이터를 분석하기에 많은 시간이 걸렸다. 이러한 문제점을 보완한 RIA(Rich Internet Application)환경의 DNA서열 분석도구의 개발은 생물학자들의 편리성과 생산성을 제고할 것으로 보인다.

본 연구에서 제시한 DNA서열분석 도구는 GenBank 파일 검색 및 분석으로 국한되어 이루어졌다. 실제 DNA서열분석도구의 구현은 아주 복잡하고 방대한 일이어서 결코 한번에 완성되어 질 수 없으며 지속적인 피드백을 받아 업그레이드가 이루어져야 한다. 이번 연구는 DNA서열분석 도구를 웹2.0기반에서 활용 할 수 있는 초석을 제시 하였다. 향후 지속적인 업그레이드를 통해 보다 정밀하고, 신속한 분석도구 자리매김할 것으로 판단된다.

향후 연구로 단백질 데이터 뱅크를 이용하여 PDB(Protein Data Bank)를 3D로 표현하는 프로그램을 추가하여 생명정보학 연구에 더욱 효율적인 도구로 개발할 계획이다.

참고문헌

- [1] R.D. Fleischmann et al., "Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd.", *science*. Vol.269, pp.496-512, 1995.
- [2] 김상수, "유전자 기능 분석을 위한 바이오인포매틱스 기법 소개", *대한간학회지*, 제 10권, 제 1호, 11-21쪽, 2004.
- [3] 니케이 바이오테크, "최첨단 리포트 유전자 비즈니스", 김영사, 2002.
- [4] 최범순, 이경희, 권해룡, 조완섭, 이충세, 김영창, .

"웹 기반 통합 유전체 분석 시스템의 설계 및 구현", *한국멀티미디어학회논문지*, 제7권, 제3호, 408-417쪽, 2004.

- [5] Berriman, M., Rutherford, K., "Viewing and annotating sequence data with Artemis", *Bioinformatics*, Vol 4(2), 2003.
- [6] Rutherford, K., Parkhill, J., Crook, "Artemis: Sequence visualization and annotation", *Bioinformatics*, Vol.16(10), pp. 944-945, 2000.
- [7] URL: <http://www.sanger.ac.uk/Software/Artemis>
- [8] L., Catherine, "A Web interface generator for molecular biology program in Unix". *Bioinformatics*, Vol.7, No.1, pp. 73-83, 2001.
- [9] 강주영, 용환승, "웹 2.0의 개념과 전망", *정보과학회지*, 제 25권, 제 10호, 5-11쪽, 2007.
- [10] 옥상훈, "예제로 배우는 Adobe 플렉스2", 에이콘, 2006.
- [11] 이호석, "XML과 응용의 고찰", *한국멀티미디어학회지*, 제10권, 제1호, 1-13쪽, 2006.
- [12] 김태환, 정병진, 손현석, 조영화, "바이오인포매틱스 ESTs 서열분석", *생능출판사*, 30-32쪽, 2003.
- [13] 신현삼, 배경희, 박강희, "개발자를 위한 플렉스2 실무테크닉", *성안당*, 2007.
- [14] 제임스티스달, 박현석, "펄로 시작하는 바이오인포매틱스", *한빛미디어*, 2002.
- [15] 성중희, 김동규, "서열 정렬 알고리즘의 연구 동향", *한국멀티미디어학회 03 춘계학술발표논문집*, 571-574쪽, 2003.
- [16] 박현석, 정철. *자바로 배우는 바이오인포매틱스*. 사이텍미디어, 2006.

저자 소개

김 명 관(정회원)



- 1981년 3월~1985년 2월 송실대학교 전자계산학과 학사
 - 1985년 3월~1987년 2월 송실대학원 전자계산학과 석사
 - 1996년 9월~2004년 2월 송실대학원 컴퓨터학과 박사
 - 1989년 8월~1993년 2월 한국전자통신 연구소 인공지능연구실 연구원
 - 1993년 3월~2007년 2월 서울보건대학 컴퓨터정보과 부교수
 - 2007년 3월~현재 을지대학교 의료산업학부 의료진산전공 부교수
- <관심분야 : 인공지능, 자연어처리, 질의응답시스템, 시멘틱 웹>

조 충 효(준회원)

- 2007년 2월 을지대학교 컴퓨터정보과 졸업