

모바일 환경에서 BioMart 데이터베이스에 대한 생물정보데이터의 접근 및 검색

Accessing and Retrieving Bioinformatics Data From Biomart on a Mobile Device

유석종, 박준호*, 임종태*, 이지희*, 포미미*, 김미경*,
김현주*, 조미림*, 류제운**, 김학용**, 유재수*
한국과학기술정보연구원 슈퍼컴퓨팅본부,
충북대학교 정보통신공학과*, 충북대학교 생화학과**

SeokJong Yu, Junho Park*, JongTae Lim*,
JiHee Lee*, Bao WeiWei*, MiKyoung Kim*,
HyunJu Kim*, MiRim Jo*, JaeWoon Ryu*,
HakYong Kim*, JaeSoo Yoo*
KISTI Supercomputing Center,
Chungbuk National University*

요약

최근 유전자 염기서열분석방법이 새롭게 등장하면서 대량의 생물정보의 데이터가 축적되고 있다. 다양한 생물종의 유전체 정보에 대하여 관련 연구자들에게 유용한 정보를 제공하기 위해서는 분석된 유전체 정보를 빠르게 전달할 필요가 증가하고 있다. 본 연구에서는 분석된 유전체 정보에 대한 데이터베이스를 구축하고 이를 모바일 환경에서 손쉽게 탐색할 수 있는 클라이언트 앱을 개발하였다. 유전체 정보에 대한 데이터베이스는 BioMart를 사용하였다. BioMart의 데이터베이스 스키마를 기반으로 웹서비스 방법을 통하여 모바일 환경의 기기와 연동할 수 있도록 하였다. 개발된 시스템의 검증을 위해서 BioMart의 유전자내용을 검색 할 수 있는 질의를 통하여 모바일기기에서 인간의 유전체 정보를 네비게이션 할 수 있도록 하였다.

I. 서론

유전체 정보를 비롯한 생명정보가 증가함에 따라 이를 효율적으로 구축할 수 있는 다양한 도구가 개발되었다. BioMart[1]는 EBI가 개발한 데이터 웨어하우스 시스템으로 간결하며, 대규모의 데이터를 다룰 수 있고, 통합된 질의 시스템을 구성하고 있다[2]. BioMart를 이용한 대표적인 데이터베이스로는 Ensemble이 있으며, 이를 통해 다양한 종의 유전체 정보를 탐색할 수 있다.

이밖에도 척추동물의 유전체를 다루고 있는 VEGA, 단일 염기서열 다형성(SNP)에 대한 데이터베이스를 관리하는 dbSNP등이 있다. 특히 최근에 다양한 생명정보에 대한 통계분석을 수행하는 R패키지를 지원하기도 한다[3].

더욱이, 유전체 염기서열을 얻은 이후 해당 유전체정보에서 어떤 유전자와 단백질이 존재하는지를 찾아내는 과정에서 분석의 자동화를 위해 PICR, DAS그리고 BioMart를 연계하여 데이터웨어하우스를 구축하는 파이프라인이 정립되어 사용되고 있다[4].

BioMart에서는 자체적인 질의문법을 정의하고 이를

이 논문은 2011년 교육과학기술부의 지원(지역거점연구단육성사업/충북BIT연구중심대학육성사업단)과 교육과학기술부와 한국산업기술재단의 지역혁신인력양성 사업으로 수행된 연구결과임.

통해 다양한 방법을 통해 클라이언트에게 검색 결과를 제공한다.

기본적인 방식으로 웹페이지를 통해 사용자에게 관련 정보를 전달하며, 이외에도 웹서비스를 통해서도 관련된 정보를 제공한다[5].

본 연구에서는 BioMart에서 제공하는 웹서비스를 이용하여 모바일 환경에서 유전체 정보를 검색하여 탐색할 수 있도록 하였다.

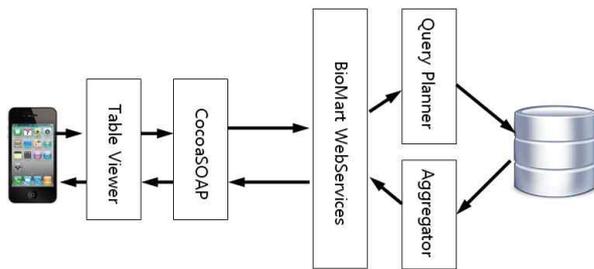
II. 시스템 설계

1. 시스템 아키텍처

BioMart는 MySQL을 기반으로 구축된 데이터 웨어하우스로 기본적인 데이터베이스 스키마는 스타 스키마로써 간결함으로 추구하고 빠른 검색수행을 구현하였다. 또한 전체 유전체 정보중에 특정내용을 빠르게 검색할 수 있도록 질의에 대한 XML정의를 통해 다양한 시스템 간의 연동성으로 보장하는 장점이 있다.

모바일 기기에서 사용자는 BioMart의 질의를 정의하여 검색을 요청하면, 모바일기기에서 BioMart웹서비스를 통해 검색요청을 서버에 전달하며, 이를 MySQL에 구축된 생명정보를 검색하여 사용자에게 전달할 수 있도록

하였다(그림1).



▶▶ 그림 1. BioMart와 연계를 위한 시스템 설계

2. 시스템 구현

사용된 모바일기기는 아이폰의 iOS4를 이용하였으며, Objective-C를 이용하여 개발하였다. 특히 웹서비스를 이용한 검색은 iOS4의 Cocoa Frameworks를 사용하였다[3]. 시스템개발은 Xcode4를 활용하였으며, 아이폰과 아이패드의 환경에서 실행할 수 있도록 구현하였다.

Ⅲ. 구현 결과

모바일기기는 작은 화면을 통해 웹페이지를 접근해야 하므로 기존 구축된 BioMart 웹 사이트를 이용하는데 불편하다. 특히 다양한 유전체 정보를 검색하고 이를 연구에 활용하기 위해 모바일 환경을 사용할 경우에 이를 지원하기 위한 적절한 사용자 인터페이스와 조작환경을 지원해야 한다.

본 연구에서 개발된 BioMart 데이터베이스에 대한 모바일 환경은 웹페이지에서 조작하기 힘든 질의 생성을 모바일 환경에서 손쉽게 생성하고 검색하며, 이를 자신의 모바일 기기에서 저장 관리할 수 있도록 하였다.



▶▶ 그림 2. 사용자 질의에 대한 검색 결과화면

그림 2는 사용자가 원하는 질의를 생성하여, 이를 통해 검색된 결과를 모바일 환경에 적합한 테이블 형태의 결과로 확인하는 화면으로 다양한 검색을 수행할 수 있다.

사용자가 검색한 결과는 휴대기기에 저장하여, 사용자가 추후에도 지속적으로 내용을 확인 할 수 있도록 하였으며, 이를 기반으로 추가적인 외부 데이터베이스 정보를 접근할 수 있도록 하였다. 사용자는 자신의 관심 있는 검색의 결과에서 이에 해당하는 추가적인 외부 데이터정보를 통합적으로 네비게이션 할 수 있어 보다 연구에 효과적으로 활용할 수 있을 것이다.

IV. 결론

본 연구에서는 최근 확산되고 있는 스마트 폰 모바일 환경에서 유전체 정보를 검색하고 이를 손쉽게 관리할 수 있는 시스템을 설계하고 구현하였다. 특히 생명과학 분야의 정보는 구조가 복잡하고 다양한 종류의 정보들이 종합되는 만큼 이와 같은 편리한 모바일 작업환경은 향후 연구의 효율성과 관리의 편리성을 제공할 것이며, 특히 생명과학분야 연구 활동이 발생하는 실험실에 적합한 연구 환경을 제공할 수 있을 것이다.

향후 생명과학분야는 대규모의 복잡한 생물정보데이터를 효율적으로 검색하고 이를 통해 새로운 연구 활동이 신속히 이루어질 수 있는 연구 환경에 모바일 장비는 중요한 기능을 수행할 것으로, 바이오분야 연구자들을 위한 지속적인 기능개선을 수행할 예정이다.

■ 참고 문헌 ■

- [1] Biomart, <http://www.biomart.org/>
- [2] Haider S, Ballester B, Smedley D, Zhang J, Rice P, Kasprzyk A. "Biomart Central Portal - unified access to biological data." *Nucleic Acids Res.* 37, pp. W23-W27, 2009.
- [3] Durinck S, Moreau Y, Kasprzyk A, Davis S, De Moor B, Brazma A, Huber W. "BioMart and bioconductor: a powerful link between biological databases and microarray data analysis." *Bioinformatics* 21, pp. 3439-3440, 2005.
- [4] Jones P. "Analysing proteomics identifications in the context of functional and structural protein annotation: integrating annotation using PICR, DAS, and BioMart." 696, pp. 107-121, 2011.
- [5] <http://developer.apple.com/devcenter/ios/index.action>