

미생물 다양성 분석을 위한 웹기반의 생물정보도구 개발

Web-based Research Assistant Tools for Analysis of Microbial Diversity

강병철^{*,**}, 김현진^{**}, 박준형^{***}, 박희경[§], 김철민^{*,**,*§}

Byeong-Chul Kang^{*,**}, Hyun-Jin Kim^{**}, Jun-Hyung Park^{***},
Hee-Kyung Park[§], and Cheol-Min Kim^{*,**,*§}

* 동서대학교 응용생명공학부

** 부산대학교 대학원 생물정보학협동과정

*** 부산대학교 의과대학부설 부산지능센터

§ 부산대학교 의과대학 생화학교실

요 약

생태학, 환경공학, 임상진단 등 생물학 분야에서 미생물의 다양성 연구의 중요성이 대두되고 그 연구가 집중하고 있다. 특히 16S rRNA를 분자지표로한 DNA 염기서열 분석방법이 널리 사용되고 있다. 본 논문에서는 16S rRNA의 염기서열 분석과정을 각 단계별로 자동화하고, 생물학자들의 결과 판단이나 사용상의 편의를 도모하기 위하여 웹기반의 미생물 다양성 분석 어플리케이션을 개발하였다. 이를 위하여 단계별 자동화 및 인터페이스 개발에 적합한 폴더-프로세스-필터 모델을 고안하고 적용하였다. 제공되는 생물정보분석도구는 서열입력, 서열방향교정, 다중서열정렬 및 가시화, 서열동정 등의 분석이 있으며, 각 결과는 계통분류도구와 호환가능하도록 하였다. 또한 신생아의 장내 세균총에 대한 분석을 수행하여 개발된 도구의 유용성을 확인하였다. 개발된 웹 어플리케이션은 리눅스 시스템 상에서 Perl 과 CGI를 이용하였으며, <http://home.pusan.ac.kr/~genome/tools/rat.htm>으로 접속하여 사용할 수 있다.

Abstract

The study of available genotypes (biodiversity analysis) in bacterial communities is of growing importance in several fields such as ecology, environmental technology, clinical diagnostics, etc. These culture-independent genotyping techniques, especially amplifying 16S rRNA genes, attempt to overcome some shortcomings of conventional cultivation method. Biodiversity analysis based on molecular technique were laborious for base-calling chromatogram, trimming primer sites, correcting strand directions, electing representative operation taxonomic units (OTU), etc. Also, biologists wanted intuitively to confirm results of the above processes. For making up these demands, we developed the web application based on *Folder-Process-Filter* (FPF) modeling with correspondence to classical Model-View-Controller model. The model of web application leads to keep virtues of simplicity and directness for development and management of the stepwise web interfaces. The web application was developed in Perl and CGI on Linux workstation. It can be freely accessed from <http://home.pusan.ac.kr/~genome/tools/rat.htm>.

Key words : 생물다양성분석, 생물정보학, 핵산염기서열분석, 웹어플리케이션 객체모델링, 객체지향모델링

1. 서 론

미생물 군집에서의 유전자형에 대한 연구는 생태학, 환경공학, 임상진단 등의 몇몇 분야에서 그 중요성이 대두되고 있다[1]. 종래의 방법으로는 미생물의 군집을 직접 배양하여 형태학적, 생화학적 분류법으로 그 다양성을 분석하였다. 그

러나 순수배양의 어려움, 세균의 작은 크기, 형태학적 구분의 제한성, 자동화의 어려움으로 대규모의 연구에는 많은 한계가 있다. 근래에는 이러한 전통적인 방법의 단점을 극복하기 위해서 미생물을 직접 배양하지 않고 분석하는 분자생물학적 방법들이 많이 제안되고 있다. 대표적인 분석방법들로는 전체 염색체 DNA에 대한 PFGE(Pulsed Field Gel Electrophoresis), Southern blotting 과 RFLP(Restriction Fragment Length Polymorphism), PCR-based locus specific RFLP, REP(Repetitive Extragenic Palindromic) PCR, CFLP(Cleavase Fragment Length Polymorphism), AFLP(Amplified Fragment Length Polymorphism) 시험 그리고 염기서열분석(DNA Sequencing) 등이 있다[1,2].

대표적인 분자지표인 16S rRNA는 단백질 합성에 핵심적

접수일자 : 2004년 3월 31일

완료일자 : 2004년 6월 1일

감사의 글 : 본 연구는 과학기술부 과학기술정책연구사업 기술개발용역과제(부산-0102)에 의해 지원 받았습니다.

인 역할을 하는 리보솜 유전자 중의 하나이다. 16S rRNA는 종과 속간의 분화에 따른 다형성 지역이 있어서 특정 분류군에만 존재하는 염기서열을 가지고 있다(그림 1(a)의 A부분). 동시에 이 유전자는 생명현상 유지에 필수적인 이차구조를 가지는 서열 부위가 있기 때문에 대부분의 생명체에 공통적으로 보존되어 있다(그림 1(a)의 B부분). 이러한 특성으로 인하여 16S rRNA를 분자지표로 하여 염기서열을 결정하고 그림 1(b)와 같이 계통분류학적으로 분석하면 다양한 분류군의 상호비교가 가능하다[3].

16S rRNA 서열분석을 기초한 생물다양성 분석에는 염기서열의 결정, 프라이머(primer) 부위 제거, DNA 서열의 방향 교정, 대표 OTU(operation taxonomic units) 선정, 키메라(chimera) 서열 검사, 서열동정, 다중서열정렬(multiple alignment), 유전적 거리(genetic distance) 계산, 계통수(phylogenetic tree) 작성 및 통계적 검증 등 다양한 단계의 생물정보학적 분석이 필요하다. 각각의 과정에 대한 공개 소프트웨어들이 여러 연구자들에 의해서 제공되고 있지만, 사용자 인터페이스가 일관되지 않으며, 일반 생물학자에게는 익숙하지 않은 유닉스(UNIX) 환경으로 제공되는 경우가 많다. 윈도 환경의 상용 소프트웨어가 있지만, 고가이거나 일괄처리를 하지 못하는 경우가 많다. 온라인 상에서 이를 지원하는 대표적인 데이터베이스는 RDP-II (Ribosomal Database Project-II)로서 95,000건 이상의 미생물의 16S rRNA 서열과 인간을 포함한 다수의 진핵생물의 리보솜 서열을 제공하고 있다[4]. 동시에 웹 기반의 정렬 프로그램과 PHYLIP[5] 인터페이스를 제공하여 본격적인 계통분류를 수행할 수 있도록 하였다. 그러나, 자동화가 절실한 염기서열의 전처리(다중서열정렬 이전까지의 처리과정)에 대한 개발이 미비한 실정이다.

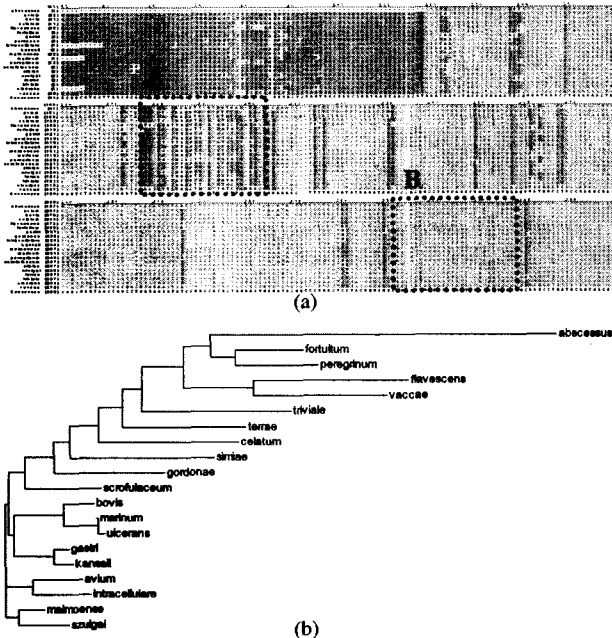


그림 1. 마이코박테리아 속내 여러종의 16S rRNA 다중서열정렬(a)과 계통분류도(b)의 예시.
Fig. 1. Exmples of (a) multiple alignment and (b) phylogenetic tree for 16S rRNA sequences of mycobacteria species.

본 논문에서는 서열 처리의 자동화와 사용자 요구에 의한 단계별(stepwise) 인터페이스 개발을 위해서 폴다-프로세스-필터(FPF) 모델을 고안하고 적용하였다. 개발된 웹기반의 생물정보분석도구는 서열입력, 서열방향교정, 다중서열정렬 및 가시화, 서열동정 등의 분석기능을 내장하고 있으며, 그 결과는 손쉽게 RDP-II에 연계되도록 하였다. 또한 신생아의 장내 세균총에 대한 분석을 수행하여 개발된 도구의 유용성을 확인하였다.

2. 시스템 구성 및 개발방법

생물정보학 소프트웨어 중에서 가장 많은 부분은 DNA 서열이나 아미노산 서열 자료를 처리하는 것이다. 다양한 서열 처리 소프트웨어들이 유닉스 환경에서 개발되었지만 유닉스에 익숙하지 않은 사용자들이 많다. 인터넷의 보급에 따라 근래에는 웹 어플리케이션으로 제공함으로써 사용성과 유지보수의 어려움을 해소하였다. 그러나 각 개발자에 의해서 제공되는 웹 어플리케이션을 각기 다른 용어와 가시화 방법을 사용하여 통일성이 부족하고 여러 종류의 어플리케이션을 파이프라인(pipeline)으로 묶어 연동시키는데 어려움이 있다.

생물정보처리를 위한 파이프라인(pipeline) 구축은 유전자 기능 탐색 분야에 많은 선행연구가 있다. PEDANT는 유전체 규모의 서열을 입력하면, 유전자 부위를 찾고 그 기능을 분석하는 전과정을 자동화하고 웹 인터페이스를 통해서 결과를 확인할 수 있도록 하였다[6]. 이는 대용량의 데이터 처리에 적합하지만, 분석과정의 단계별 확인이 필요한 분석업무에는 부적절하며 복잡한 시스템 구성으로 유지보수에 어려움이 있다. Bioperl[7]의 세부 프로젝트의 하나인 Biopipe (Bioperl pipeline framework)는 매우 유연성이 높은 방법으로 알려져 있다[8]. 그러나 Biopipe도 PEDANT와 같이 많은 분석 절차의 자동화를 위한 구조를 채택하여 각 단계에 대한 인터페이스 구현에 어려움이 있다.

따라서 생물정보처리 파이프라인을 구축하고 일관성 있는 사용자 인터페이스를 모델링할 수 있는 방법이 필요하다. 또한 모델링된 생물정보 파이프라인과 사용자 인터페이스는 웹 어플리케이션 개발시에 많이 도입되는 MVC(Model-View-Controller) 모델과 연계되도록 하여 웹으로 구현하기 쉽도록 하여야 한다.

2.1 서열분석 파이프라인의 모델링

서열 자료 처리 측면에서는 (특히, 생물다양성 분석) 각 프로세스로 간주할 수 있는 소프트웨어가 연결된다는 점에서 파이프라인으로 간주할 수 있다. 하지만 최초의 입력데이터인 염기서열자료가 각 프로세스 결과에 의해서 가감되지만 계속 연계된다는 특성이 있다. 예를 들어 그림 2(a)와 같이 다수의 염기서열 자료로부터 유사성을 확인하여 OTU를 선정하고 각 OTU의 종을 동정하기 위해서 BLAST를 수행하는 파이프라인을 고려하자. 여기서 최초의 정보로 분자지표의 서열 자료를 ClustalW에 입력하면 설정된 파라미터에 의해서 다중서열정렬을 한다. 다중서열정렬 결과에서 필터링 조건인 역치값 보다 높은 % identity를 가지는 그룹은 하나의 OTU로 묶는다. 이 때 원시 자료인 입력서열중에서 대표를 선정하던가 consensus 서열을 사용한다. 이후에 선정된 각 OTU 서열을 BLAST 검색을 수행하여 종 동정을 수행하고 원시 파일에 동정된 결과를 기록한다.

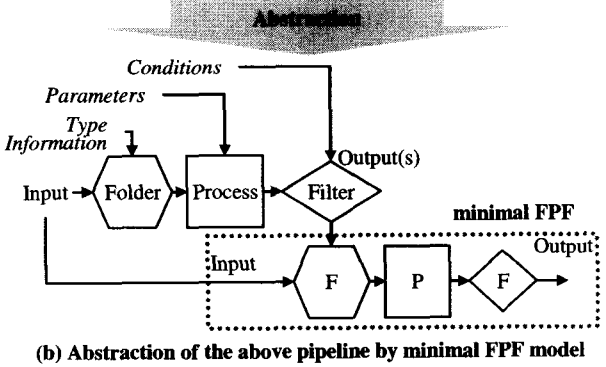
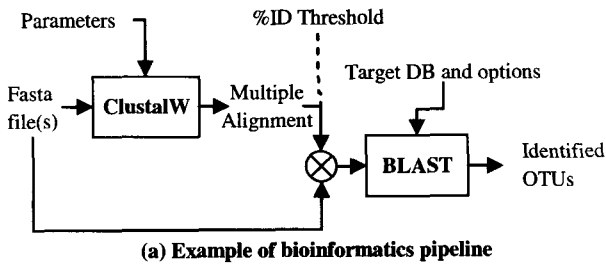


그림 2. 생물정보파이프라인(a)과 폴더-프로세스-필터 모델에 의한 추상화(b)의 예시.

Fig. 2. (a) Example of bioinformatics process pipeline and (b) its abstraction model by Folder-Process-Filter modeling.

이와 같이 서열 처리 과정에서는 원시 입력 자료를 파이프라인에서 지속적으로 사용해야 되는 경우가 많으므로 프로세스와 그 전·후 처리부분을 그림 2(b)와 같이 추상화시켰다. 폴더(folder) 클래스는 입력 자료의 저장소 역할을 하고 입력 자료의 형식 검사를 수행한다. 프로세스(process)는 이미 개발된 생물정보처리 프로그램을 캡슐화시키거나 독자적인 알고리즘이 구현되는 부분이다. 필터(filter)는 프로세스의 결과를 후반 처리하는 부분으로 결과를 조건에 따라 분류하거나 다음 처리를 위한 형식 변환 등이 주요한 역할이다. 이렇게 생물정보분석 과정은 그림 2(b)와 같이 하나의 폴더, 프로세스, 필터 객체를 가지도록 최소FPF(minimal FPF)로 모델링하고, 일련의 생물정보분석 과정은 FPF 네트워크(FPF network)로 대체시킬 수 있다.

표 1. FPF 네트워크와 유한상태모형과의 관계.
Table 1. Analogy between finite state machine model and FPF network.

Elements of Finite State Machine	Elements of FPF Network
State	States of a minimal FPF; possible all states in web application
Transaction	Changes of state by stimulated to web page
Input	User stimulus to induce state change (input of field, selection of menu, mouse click, etc.)
Output	Output results by transaction

최소FPF와 그 네트워크의 구성은 XML로 된 기술규약에 의해서 정의되도록 하였다. 즉, 단계별 생물정보분석 파이프라인은 최소FPF의 연결에 의해서 기술되도록 하였다. 기술된 FPF 네트워크는 표 1과 같이 유한상태모형(finite state machine)에 맵핑시킬 수 있다. 이와 같이 FPF 네트워크를 기술하는 XML 규격을 FNDL (FPF Network Discription Language)로 정의하였다. 향후 상태모형의 검사를 통해서 코딩을 하기전에 웹 어플리케이션 설계의 무결성을 확인할 수 있도록 할 것이다.

2.2 MVC 모델과의 연계

MVC 모델은 웹 어플리케이션을 포함한 많은 GUI (graphic user interface) 프로그래밍에서 전통적으로 사용된 모델링 기법이다. 본 연구에서는 제안된 FPF 모델에 의해서 구성된 생물정보처리과정을 웹 어플리케이션으로 구현하기 위해 프레임워크를 그림 3과 같이 설계하였다. 컨트롤러(controller)는 CGI[9]로 상속받은 StepwiseApp 클래스와 각 페이지간의 연결 정보를 관리하는 PageMode 클래스에 의해서 구현되고 뷰(view)는 TTK2 (template tool kit 2, [10])를 사용하여 사용자 인터페이스 프로그래밍과 생물정보 프로그래밍 작업이 분리되도록 하였다. 모델(model)은 ControllerModel 인터페이스를 통해서 구현되도록 하고 FPF 네트워크의 정보 대부분을 관리한다.

2.3 웹 어플리케이션 개발을 위한 코드 생성

웹 어플리케이션의 기본 코드는 XML에 의해서 정의된 FPF 네트워크 정보로부터 자동으로 생성된다. 즉, 사용자 요구에 따른 생물분석도구를 선정하고 기본적인 파이프라인의 구성을 FPF 네트워크로 표현한다. 고안된 프레임워크는 FPF 네트워크 정보에 따라 CGI로부터 상속된 코드를 생성한다. 생성된 코드에는 CGI의 기본 제어 구조, 웹 입출력을 위한 템플릿 객체, 데이터 저장을 위한 객체 정보를 포함하고 있다. 따라서, 템플릿을 이용한 웹 프로그래밍과 펄(perl)이나 외부 프로그램을 이용한 생물정보처리 프로그래밍이 구분되어진다.

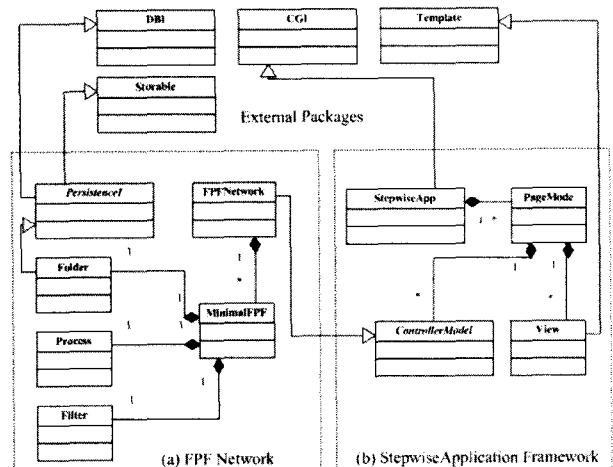


그림 3. 웹 어플리케이션 프레임워크와 FPF 네트워크의 UML 모델 다이어그램. (a) FPF 네트워크 모델 다이어그램, (b) 웹 어플리케이션 프레임워크.

Fig. 3. (a) UML Diagram of the proposed FPF network and (b) web application framework.

3. 적용 사례 결과 및 고찰

본장에서는 FPF 모델링에 의한 생물다양성 분석 웹 어플리케이션을 활용하여 실제 실험에 적용하여 그 유용성을 확인하고자 한다. 예시하는 실험의 목적은 면역계 형성에 중요한 것으로 알려진 신생아의 장내세균의 군집 다양성을 확인하는 것이다.

제안한 FPF 모델링의 유용성을 확인하기 위해 분자지표를 이용한 생물다양성 분석과정 중에서 그림 4와 같이 핵심적인 처리 과정이면서 자동화 처리가 미진한 파이프라인을 모델링하였다.

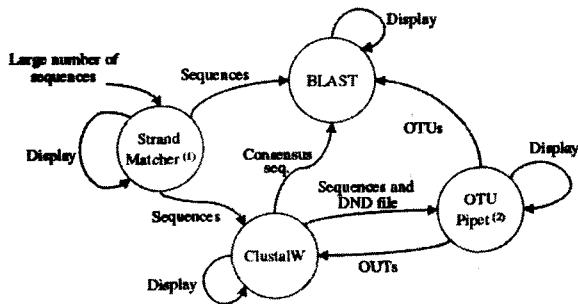


그림 4. 적용한 웹기반의 생물정보분석도구의 FPF 네트워킹의 유한상태모형도.

Fig. 4. Finite state machine diagram of case study for the proposed FPF modeling.

3.1 웹 어플리케이션의 구현

미생물의 다양성 분석시에 필요한 전처리 과정을 그림 4은 같이 FPF망으로 설계하고 프레임워크를 이용하여 기본 코드를 생성하였다. 서열의 동정을 위해서 BLAST[11]를 이용하였고 기초적인 계통분류를 위해서 ClustalW[12]를 활용하였다. 각 서열의 방향 조정과 OTU선정을 위해서는 자체 개발한 알고리즘을 구현하였다. 다중서열정렬 결과는 MVIEW[13]에 의해 보여지고 이것의 계층구조는 SVG[14]형식의 그래프파일로 제공되도록 구현하였다. 그림 4에서 (1)은 자체 개발된 도구로 *Escherichia coli* 16S rRNA를 기준으로 입력서열의 DNA방향을 교정하는 도구이다. (2)는 기준 역치값보다 높은 유사성을 보이는 서열은 하나의 서열로 간주하여 대표 OTU를 선정하는 도구이다.

개발된 각 웹 어플리케이션의 실제 화면을 그림 5에 보이며, <http://home.pusan.ac.kr/genome/tools/rat.htm>으로 접속하여 사용할 수 있다.

3.2 실험사례 : 신생아 장내 세균총의 분포 변화

구현된 웹기반의 생물정보도구의 유용성을 확인하기 위하여 실제 실험 데이터를 분석하였다. 실험 목적은 신생아의 장내 세균 분포를 확인하는 것으로 장내세균총은 인체의 건강 유지와 질병에 중요한 역할을 담당하고 있다. 특히 주산기(임신 28주부터 생후 1주) 장내세균총의 상태는 면역계에 큰 영향을 준다고 알려져 있다. 그러나 장내세균총의 상태를 알기 위한 기존의 연구들은 배양에 의해 얻어지는 세균 종만을 대상으로 하였으므로 장내세균총 전체의 상태를 반영하고 있다고 할 수 없다. 본 연구는 주산기 장내세균총의 정착과 종의 변화를 이해하기 위하여 분자생물학적 방법으로 장내균총이 형성되는 과정과 균총의 다양성을 조사하였다.

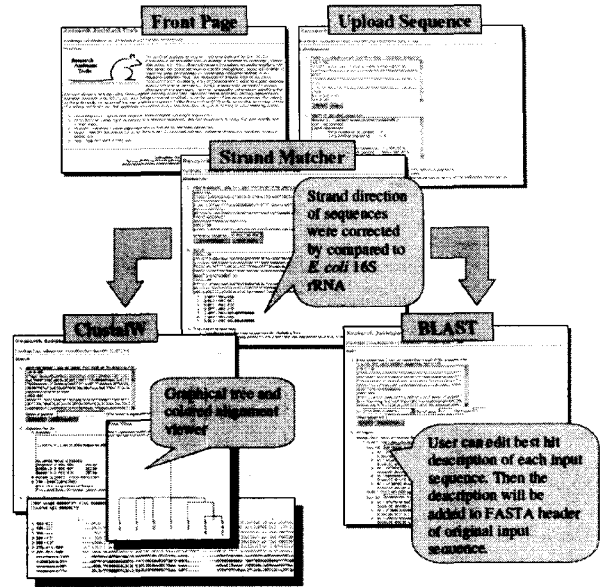


그림 5. 개발된 웹 기반의 생물다양성 분석 도구.
Fig. 5. Screen shot of the implemented web-based biodiversity analysis tools.

뒤 중합효소연쇄반응에 의해 16S rDNA 유전자를 선택적으로 증폭하여 클론화 라이브러리 구축 뒤 각 클론의 염기서열분석을 통하여 장내 균총의 다양성을 조사하였다.

그림 6은 신생아로부터 1일, 3일, 6일째에 채취된 분변으로부터 분석한 세균총의 결과이다. 중합효소반응으로 증폭시켜 클로닝한 325개의 클론중 220개는 알려진 미생물이었으며 나머지 105개는 현재까지 동정되지 않은 미생물(unidentified bacteria)로 밝혀졌다.

그림 7과 같이 동정된 세균의 변화를 보면 출생 제1일에, *Enterobacter*, *Lactococcus lactis*, *Leuconostoc citrem*와 *Streptococcus mitis*가 균총을 형성하고 있었으며 그 중 *Lactococcus lactis*가 우세균종이었다. 제3일에는 *Enterobacter*, *Enterococcus faecalis*, *Escherichia coli*, *Streptococcus mitis*와 *Streptococcus salivarius*로 균총이 형성되어 있었고 *Streptococcus salivarius*가 우세균종이었다.

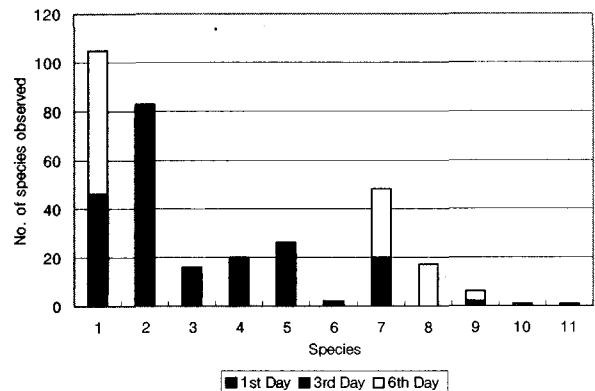


그림 6. 신생아의 장내세균총의 종 다양성.
Fig. 6. Species distribution of bacterial community of infant gut at 1st, 3rd, and 6th day from birth.

5. 참고 문헌

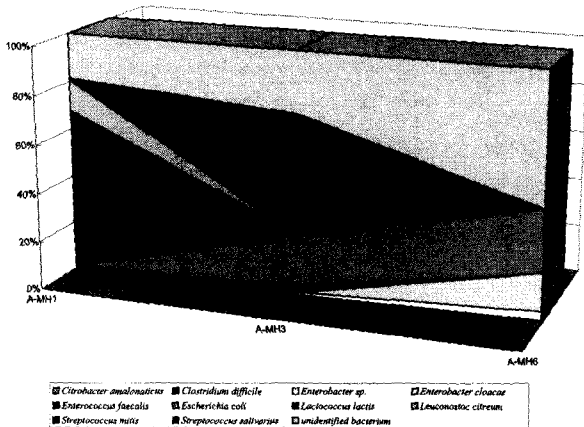


그림 7. 장내세균총의 변화 양상.
Fig. 7. Changes of the bacterial community.

제 6일에는 *Citrobacter* sp., *Clostridium difficile*, *Enterobacter* sp., *Enterobacter cloacae*와 *Escherichia coli*가 균총을 형성하였으며 우세균종은 *Escherichia coli*이었다. 이상의 동정된 균종들은 모두 배양 가능한 균종들이었으나 동정되지 않은 균종이 제 1일에 18.5%, 제 3일에 25.3%, 제 6일에 53.6%로서 기존의 배양에 의해 밝혀내지 못한 균종이 장내 세균총 형성에 매우 중요한 것으로 밝혀졌다. 이들 동정되지 않은 세균들 대부분이 배양되지 않는 미확인 세균 일 것으로 사료된다.

4. 결론 및 향후 연구

지금까지 우리는 16S rRNA 서열을 바탕으로 미생물의 다양성 분석을 지원하기 위한 웹 애플리케이션과 단계별 웹 애플리케이션 개발을 위한 모델을 제시하였다. 그러나 제안된 FPF 모델은 단계별 웹 인터페이스의 개발과 유지보수를 용이하게 할 것으로 기대한다. 또한 미생물 다양성 분석용 웹 도구는 실제 실험데이터를 분석하여 그 유용성을 확인하였다. 본 논문에서 제안하는 모델링은 유연성과 사용성을 중심으로 고안하여 다음과 같은 목적을 달성하였다. 다양한 생물정보분석 도구(서열분석도구 중점적으로)의 자동화를 제공하는 모델링을 고안하였다. 일관된 인터페이스와 활용도 낮은 파라미터 설정을 배제한 웹 기반의 사용자 인터페이스 개발을 위한 모델링 방법을 제안하였다. 그리고, 입력데이터가 분석 결과에 따라 분류되어 다른 분석 프로그램에 재입력이 가능하도록 모델링하였으며 원활한 유지보수를 위해서 웹 인터페이스 코드는 전체 애플리케이션 코드와 분리되어 개발 가능하도록 하였다. 그러나 제안된 본격적인 계통분류 패키지를 포함하고 있지 않으므로 PHYLIP, ARB[15], RDP(Ribosomal Database Project)와 연계하여 활용하면 좋을 것으로 기대한다. 향후 개발은 프로세스 관리모듈(process manager)을 추가하여 시간이 많이 소요되는 분석과정의 자동화를 가능하게 하고 PHYLIP 인터페이스를 포함하여 엄밀한 계통분류가 가능하도록 하겠다.

[1] I. Dahllof, "Molecular community analysis of microbial diversity," *Curr. Opin. Biotechnol.*, Vol. 13, pp. 213-217, 2002.

[2] D.M Olive, and P. Bean, "Principles and Applications of Methods for DNA-Based Typing of Microbial Organisms," *J. of Clin. Microbiol.*, Vol. 37, pp. 1661-1669, 1999.

[3] B. Pace, and L.L Campbell, "Homology of ribosomal ribonucleic acid of diverse bacterial species with *Escherichia coli* and *Bacillus stearothermophilus*," *J. Bacteriol.*, Vol. 107, pp. 543-547, 1971.

[4] J.R. Cole, B. Chai, T.L. Marsh, R.J. Farris, Q. Wang, S.A. Kulam, S. Chandra, D.M. McGarrell, T.M. Schmidt, G.M. Garrity, and J.M. Tiedje, "The Ribosomal Database Project (RDP-II): previewing a new autoaligner that allows regular updates and the new prokaryotic taxonomy," *Nucleic Acids Res.*, Vol. 31, pp. 442-443, 2003.

[5] J. Felsenstein, PHYLIP (Phylogeny Inference Package) version 3.5c. Distributed by the author, Department of Genetics, University of Washington, Seattle, 1993.

[6] D. Frishman, K. Albermann, J. Hani, K. Heumann, A. Metanowski, A. Zollner, H.W. Mewes, "Functional and structural genomics using PEDANT," *Bioinformatics*, Vol. 17, pp. 44-57, 2001.

[7] J.E. Stajich, D. Block, K. Boulez, S.E. Brenner, S.A. Chervitz, C. Dagdigan, G. Fuellen, J.G.R. Gilbert, I. Korf, H. Lapp, H. Lehvaslaiho, C. Matsalla, C.J. Mungall, B.I. Osborne, M.R. Pocock, P. Schattner, M. Senger, L.D. Stein, E.D. Stupka, M. Wilkinson, and E. Birney, "The Bioperl Toolkit: Perl modules for the life sciences," *Genome Research*, Vol. 12, pp. 1161-1168, 2002.

[8] C.A. Letondal, "Web interface generator for molecular biology programs in Unix," *Bioinformatics.*, Vol. 17, pp. 73-82, 2001.

[9] J. Erlbaum, <http://search.cpan.org/~markstos/CGI-Application-3.22/lib/CGI/Application.pm>

[10] A. Wardley, <http://search.cpan.org/~abw/Template-Toolkit-2.13/lib/Template.pm>

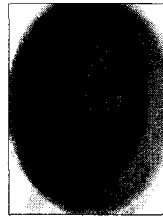
[11] S.F. Altschul, W. Gish, W. Miller, E.W. Myers, and D.J. Lipman, "A basic local alignment search tool," *Journal of Molecular Biology*, Vol. 215, pp. 403-410, 1990.

[12] J.D. Thompson, D.G. Higgins, and T.J. Gibson, "CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice," *Nucleic Acids Res.*, Vol. 22, pp. 4673, 1994.

[13] N.P. Brown, C. Leroy, and C. Sander, "MView: A Web compatible database search or multiple alignment viewer," *Bioinformatics*, Vol. 14, pp.

380-381, 1998.

[14] SVG, <http://www.adobe.com/svg/main.html>
 [15] W. Ludwig, O. Strunk, R. Westram, L. Richter, H. Meier, Yadhukumar, A. Buchner, T. Lai, S. Steppi, G. Jobb, W. Förster, I. Brettske, S. Gerber, A.W. Ginhart, O. Gross, S. Grumann, S. Hermann, R. Jost, A. König, T. Liss, R. Lüßmann, M. May, B. Nonhoff, B. Reichel, R. Strehlow, A. Stamatakis, N. Stuckmann, A. Vilbig, M. Lenke, T. Ludwig, A. Bode, and K-H. Schleifer, "ARB: a software environment for sequence data," *Nucleic Acids Res.*, Vol. 32(4), pp. 1363-1371, 2004.



박희경(Hee-Kyung Park)

1991년 : 인하대학교 대학원 화학공학과
 생물공학전공(공학석사)
 1997년 : 동 대학원 화학공학과
 생물공학전공(공학박사)
 1999년~2003년 : (주)에스제이하이테크
 생명화학연구소 연구소장
 2003~현재 : 부산대학교 의과대학 생화학
 교실 박사후 연구과정

관심분야 : DNA chip, Protein chip, Genomics
 Phone : 051-246-4828
 FAX : 051-248-1118
 E-mail : dnachiphk@hotmail.com

저 자 소 개



강병철(Byeong-Chul Kang)

1996년 : 부경대학교 기계공학과(공학사)
 1998년 : 동 대학원 기계공학과
 자동제어전공(공학석사)
 2004년 : 부산대학교 대학원 생물정보협동
 과정(이학박사)
 2004년~현재 : 동서대학교
 응용생명공학부 초빙교수

관심분야 : 비교유전체학, 유전체분석, 생물정보학
 Phone : 051-320-1859
 E-mail : bckang@{dongseo.ac.kr, pusan.ac.kr}

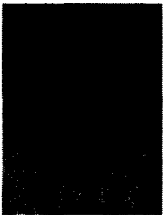


김현진(Hyun-Jin Kim)

1999년 : 신라대학교 전자계산학과
 (이학사)
 2003년~현재 : 부산대학교 대학원
 생물정보협동과정 석사과정

관심분야 : 생물정보학

Phone : 051-246-4828
 E-mail : asever@pusan.ac.kr



박준형(Jun-Hyung Park)

1998년 : 부경대학교 기계공학과(공학사)
 2000년 : 동 대학원 메카트로닉스학과
 (공학석사)
 2004년 : 부산대학교 대학원 생물정보협동
 과정 박사수료
 2004년~현재 : 부산지놈센터 연구원

관심분야 : 마이크로어레이, 유전체분석, 생물정보학
 Phone : 051-246-4828
 E-mail : jhpark98@pusan.ac.kr



김철민(Cheol-Min Kim)

1988년 : 부산대학교 의과대학(의학사)
 1990년 : 동 대학원 생화학전공(의학석사)
 1993년 : 동 대학원 생화학전공(의학박사)
 1998년~현재 : 부산대학교 의과대학
 생화학교실 부교수
 2000년~현재 : 한국유전체학회 학술지
 편집위원(생물정보학담당)

2000년~현재 : 부산대학교 생물정보협동과정 교수
 2001년~현재 : 부산대학교 의과대학부설
 부산지놈센터 센터장
 2002년~현재 : 한국생물정보학회 국제협력위원회 위원

관심분야 : DNA chip, Comparative Genomics,
 Biomedical Informatics
 Phone : 051-240-7725
 FAX : 051-248-1118
 E-mail : kimcm@pusan.ac.kr