정 동 수 박사
생물학전문연구정보센터 팀장
포항공과대학
Tel. 82-54-279-8198, Fax. 82-054-279-5540
E-mail : viroid97@bric.postech.ac.kr
http://bric.postech.ac.kr
Address : 경북 포항시 남구 효자동 산 31 <우:790-784>

◈ 연구관심분야
- 바이러스 및 바이로이드 복제메커니즘
- 생물정보학
- 생물산업 동향분석

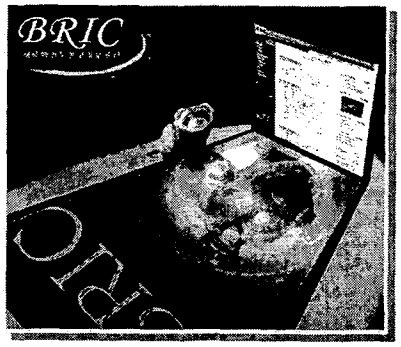
◈ 학 력
경북대학교 미생물학과 학사, 1990. 8.
경북대학교 미생물학과 석사, 1993. 2.
경북대학교 미생물학과 박사, 1998. 2.
영남대학교 미생물학과 박사후 연수 1998.3

◈ 주요경력
1. 영남대기초과학연구소 연구원
2. 일본이화학연구소 연구원
3. 생물학전문연구정보센터 팀장

◈ 연구 실적 요약
- 학술잡지 논문발표 : 10 편
- 학술컨퍼런스 논문발표 : 8 편
- 국제학회 기조연설, 초청강연, 초청세미나 : 3 회
- 저서/edited books : 1 권
- 연구과제 프로젝트 : 4 건

BRIC

Trends in Microbial Genomics and Bioinformatics

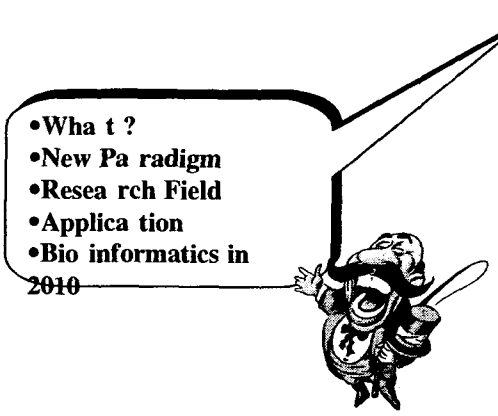


Dong Soo Jung

Biological Research Information Center (BRIC)

BRIC

Trends in Microbial Genomics and Bioinformatics



1. Introduction
2. Microbial Genome
3. Korean Bioinformatics
4. Bioinformatics Market
5. Education
6. BRIC's Bioinformatics

BRIC

What?

I. Introduction

Bio + **informat** + **ics**
Biology Information Science

It is called as ;

Bioinformatics, Biocomputing, Biomedical computing, Computational biology, Information biology, Biological data mining, *in silico* biology

- apply Mathematics/Statistics/Computer Science to Biology
- storage, analysis, interpretation of biological information
- construction of biological information infrastructure
- solving problems arising from biology using methodology from computer science.

knowledge-based discovery

BRIC

Research Fields of Bioinformatics — I. Introduction

Medical Informatics
eg., Data bases
Telemedicine

Biological engineering
eg., Microbial processing – simulation and process control

***Bio*-Chemical Informatics**
Metabolomics
Chemical-biological relationships

Sequence Analysis
Pairwise sequence alignment
Database searches
Multiple sequence alignment
Functional prediction

Structure Analysis
RNA secondary structure prediction
RNA tertiary structure modeling
Prediction of protein structural features

Genome Analysis
Comparative genomics
Gene expression analysis
Pathway databases
Path computations

Relevant computational techniques
Data processing for DNA chip or proteomic analysis
Natural language processing for biology
Simulation of artificial cell life

Knowledge Formulation
Data base formulation for biological informations
Hypothesis formation
***In silico* experimentation**
New knowledge discovery:Datamining

BRIC

New Paradigm — I. Introduction

인간유전체연구(HGP)결과 발표

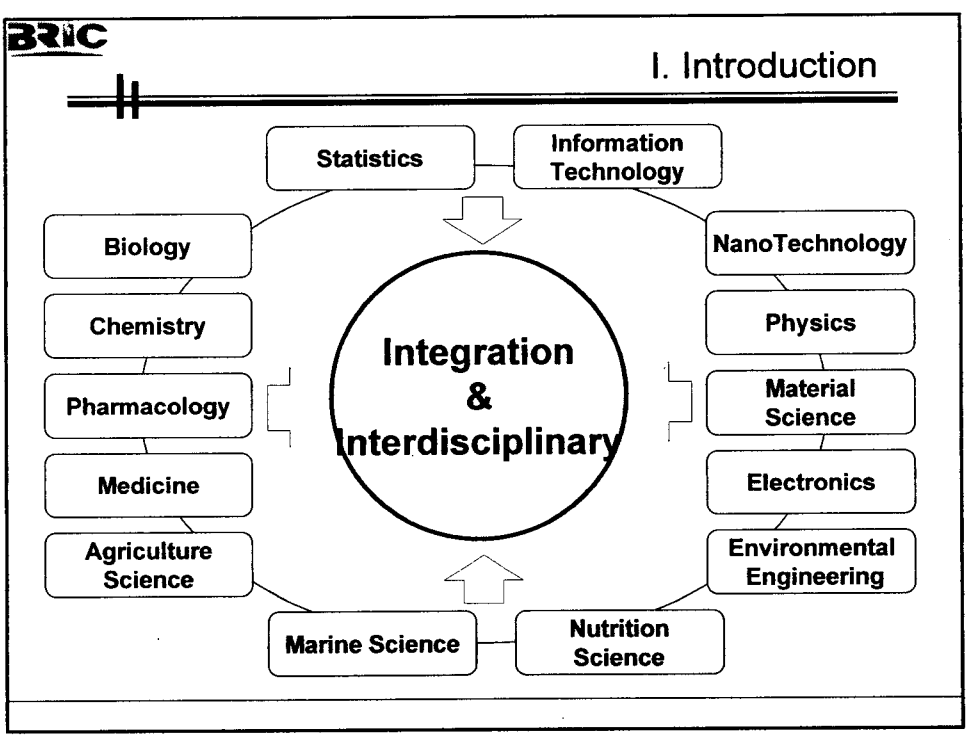
초고속 분석능력과 대용량 정보처리능력(high-throughput)

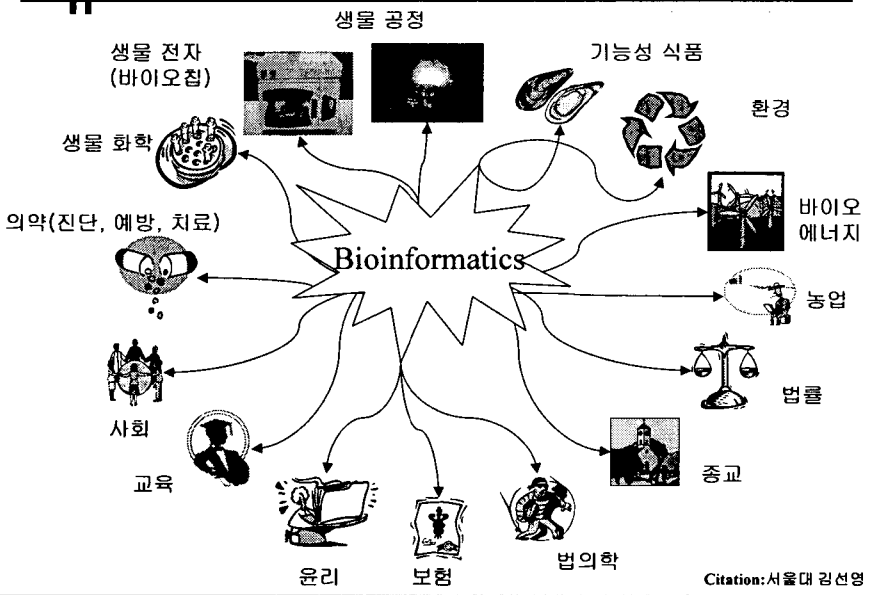
학문간의 통합과 수렴

Bioinformatics, Genomics, Proteomics

인식의 변화: 세포/조직/기관으로 구성된 유기체 ▶ 정보의 집합체
연구환경의 변화: ***in vitro, in vivo*** ⇒ ***in silico***
신약 후보물질 발굴 시간·비용 비교

성 공 확 률:	1/1만~1/10만	▶	1/10~1/100
평균개발기간:	10~15년	▶	5~7년
개 발 비 용:	100~1,200억원	▶	10~30억원





Bioinformatics in 2010

- ***In Biology***
  - ***In silico* experiment**
- ***In Healthcare***
  - **Fast drug discovery, and personalized drug.**
  - **Understand diseases at the molecular level.**
  - **Fast diagnosis**
- ***In Universities***
  - **Maybe at least one bioinformatician will be needed for biological research.**
- ***In Industry***
  - **Biotech industries will not survive without introduction of bioinformatics tools and knowledge.**
  - **Many jobs will be created.**

Trends in Microbial Genomics and Bioinformatics

1. Introduction
2. Microbial Genome
3. Korean Bioinformatics
4. Bioinformatics Market
5. Education
6. BRIC's Bioinformatics

Overview
Why Microbe?
Genome analysis
Research Institute
Website
In Silico

Microbial Genome Project--worldwide

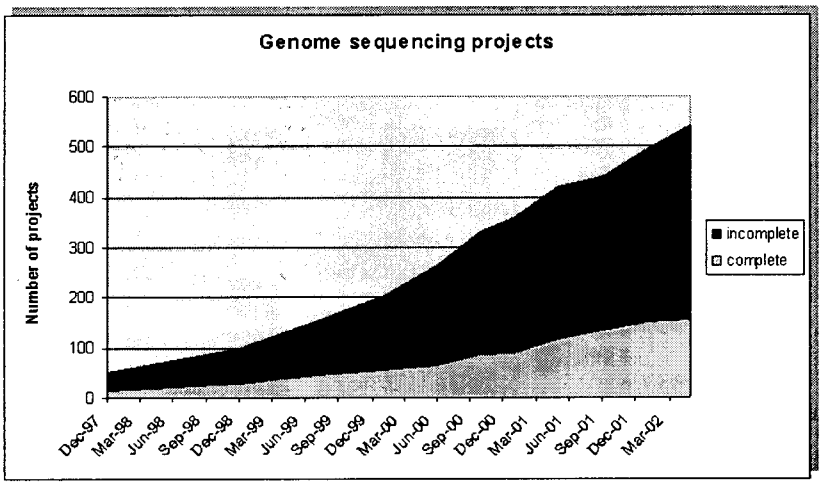
GOLD™: Genomes OnLine Database HomePage

Contact: GOLD	Last Update: June 5, 2002	Sponsored by Integrated Genomics Inc.
	Search GOLD: 570 genome projects	
Published Complete Genomes: 93	Prokaryotic Ongoing Genomes: 283	Eukaryotic Ongoing Genomes: 194

GOLD: Genomes Online Database, is a World Wide Web resource for comprehensive access to information regarding complete and ongoing genome projects around the world.

Sponsored by Integrated Genomics Inc.

Microbial Genome Project--worldwide



자료: **GOLD: Genomes Online Database**

Why Microbes?

1. Microbes make up about 60% of the Earth's biomass
2. Microbes play a critical role in natural biogeochemical cycles
3. Surviving and thriving in an amazing diversity of habitats, in extremes
4. Offer us new solutions to longstanding challenges in environmental and
5. Waste cleanup, energy production and use, medicine, industrial processes, agriculture, and other areas.
6. Biological underpinnings of climate change and the contributions of microbes to Earth's biosphere
7. Traditional commercial uses for microbes in the brewing, baking, dairy, and other industries.

BRIC

Genome Size

Organism	Haploid genome size (Mb)	Predicted number of genes
Arabidopsis thaliana (plant)	130	~25,000
Caenorhabditis elagans (worm)	100	18,424
Drosphila melangaster	180	13,601
Escherichia coli	4.7	4,288
Homo sapiens (human)	30,000	45,000 – 60,000
Saccharomyces cerevisiae (yeast)	13.5	6,241

BRIC

Steps of Genome Analysis

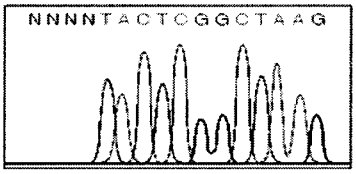
1) Genome sequence assembled
2) Identify repetitive sequences – mask out
3) Gene prediction – train a model for each genome
4) Look for EST and cDNA sequences
5) Genome annotation
6) Microarray analysis
7) Metabolic pathways and regulation
8) Protein 2D gel electrophoresis
9) Functional genomics
10) Gene location/gene map
11) Self-comparison of proteome
12) Comparative genomics
13) Identify clusters of functionally related genes
14) Evolutionary modeling

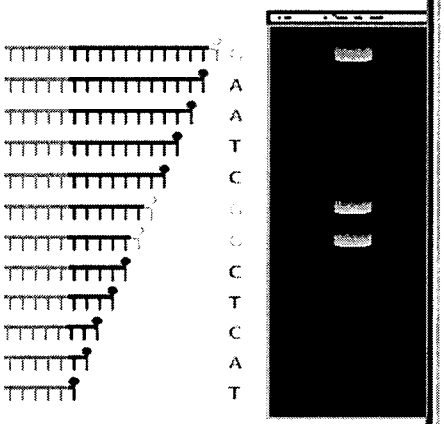
BRIC

Genome Sequencing

Cycle Sequencing

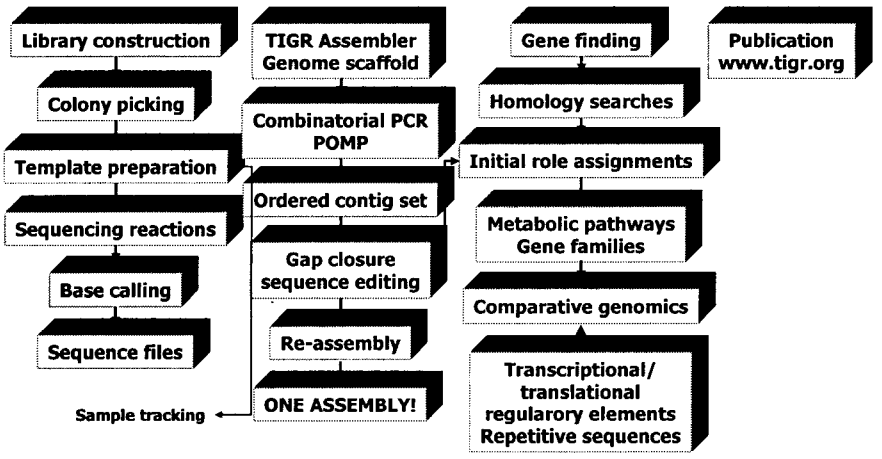
The simulated gel image is read from bottom to top, starting with the smallest fragment.

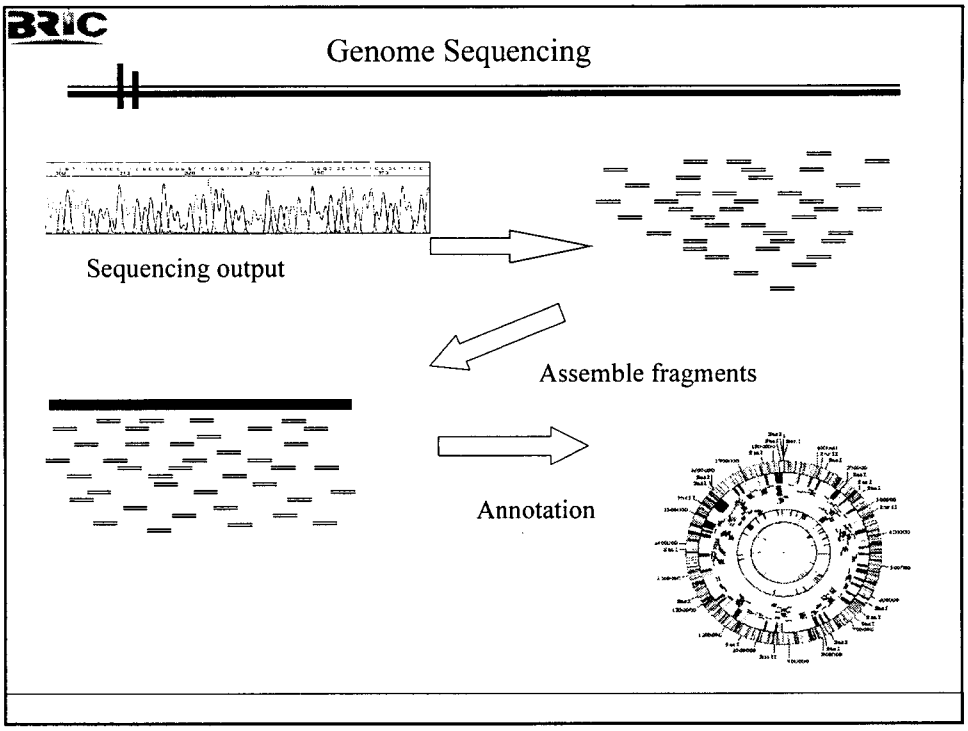


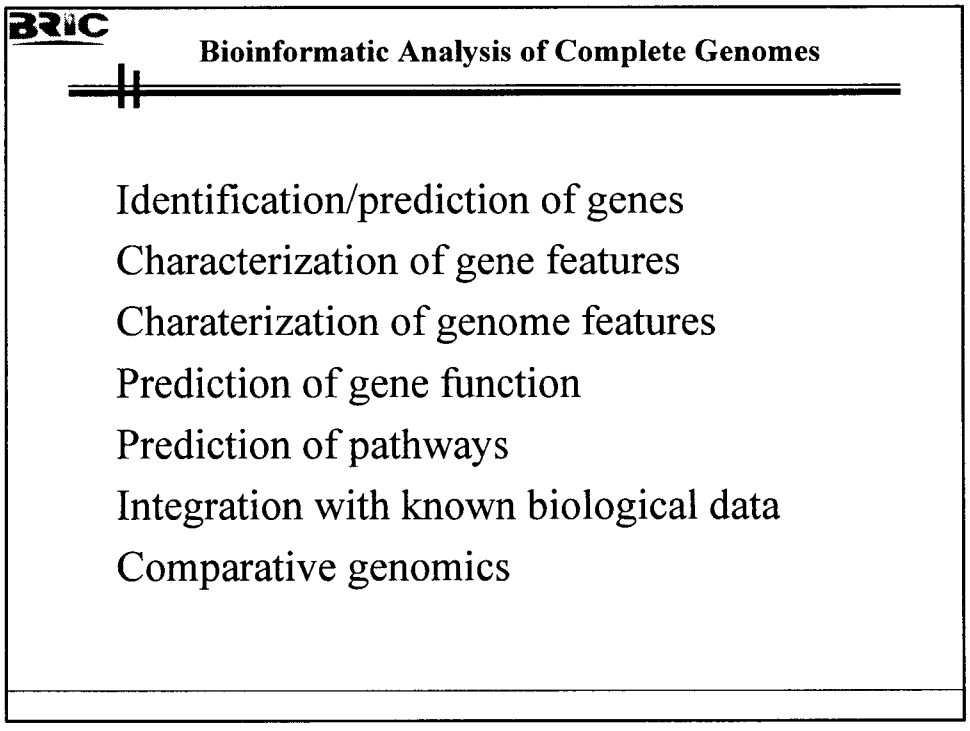


BRIC

A Microbial Genome Sequencing Project at TIGR




BRIC
Genome Sequencing
Sequencing output
Assemble fragments
Annotation


BRIC
Bioinformatic Analysis of Complete Genomes
Identification/prediction of genes
Characterization of gene features
Charaterization of genome features
Prediction of gene function
Prediction of pathways
Integration with known biological data
Comparative genomics

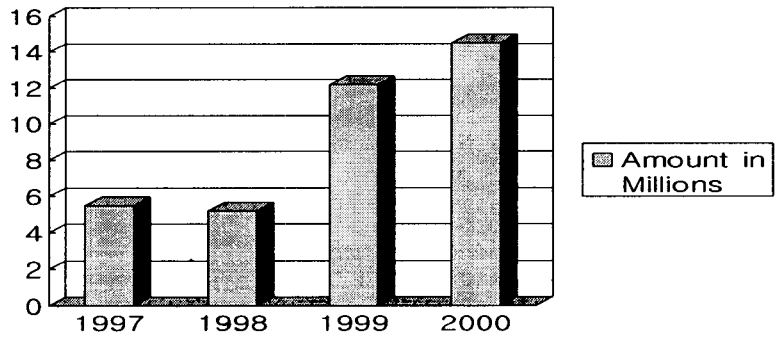
What are we learning from genome analysis?

The complete genetic blueprint for an organism:

- **Identifies novel biochemical and metabolic pathways;**
- **Identifies genes and pathways for bioremediation;**
- **Identifies organisms that have the potential to provide renewable energy sources;**
- **Identifies genes involved in adaptation to extreme environments;**
- **Identifies enzymes of industrial importance;**
- **Identifies disease-causing genes;**
- **Identifies new targets for development of antibiotics and vaccines.**

DOE Microbial Genome Program Funding

National Institute of Allergy and Infectious Diseases (NIAID)
Department of Energy (DOE),
National Science Foundation (NSF).



Totals include 1$2.4M and 2$5.8M from DOE Carbon Management Science Program (CMSP).
CMSP is part of the federal Climate Change Technology Initiative

DOE-Funded Microbial Genomes

Sequencing Microbial Genomes to Uncover Potential Applications Relevant to DOE Missions

Carbon Sequestration
-- *Chlorobium tepidum*
--*Rhodopseudomonas palustris*

Energy Production
--*Methanobacterium thermoautotrophicum*
--*Methylococcus capsulatus*

Bioremediation
--*Deinococcus radiodurans*
--*Pseudomonas putida*

Industrial Processes
--*Halobacterium halobium*
--*Pyrococcus furiosus*

Cellulose Degradation
--*Clostridium thermocellum*
--*Thermobifida fusca*

Technology Development, Pilot Projects
--*Mycoplasma genitalium*
--*Xylella fastidiosa*

National Institute of Allergy and Infectious Diseases (NIAID)

NIAID

Division of Intramural Research(four)
Division of AIDS;
Division of Allergy, Immunology and Transplantation;
Division of Microbiology and Infectious Diseases
Division of Extramural Activities
Vaccine Research Center

DMID Supported Large-Scale Genome Sequencing Projects

▶ **COMPLETED BACTERIAL GENOME PROJECTS**
: ***Chlamydia pneumoniae*** 외 16종

▶ **ONGOING SEQUENCING PROJECTS**
: ***Anopheles gambiae*** 46종

NIAID will give priority consideration to large scale sequencing projects for the following organisms: *Aedes aegypti, Anopheles gambiae, Brugia malayi, Coccidioides immitis, Group B streptococcus, Histoplasma capsulatum, Rickettsia rickettsii, Toxoplasma gondii and Trichomonas vaginalis.*

BRIC

National Institute of Allergy and Infectious Diseases (NIAID)

Goals

The Microbe Project has three broad goals:
To build needed infrastructure,
To promote research,
To develop human resources and an informed public.

Genome-enabled microbial research holds enormous promise for understanding life at its most basic level, and for enabling breakthrough applications in health, agriculture, biotechnology, the environment and national defense.

The three major components of infrastructure needed to support microbial genomics research are
1) Genome sequences
2) Tools,technologies and biological resources
3) Databases and bioinformatics

BRIC

Structural Genomics and Proteomics Technology Branch

New Branch of Division of Cell Biology and Biophysics,The National Institute of General Medical Sciences (NIGMS)

Purpose: support studies that take a genomics or computational approach to determining protein structures and functions

Berkeley Structural Genomics Center
Focus on two bacteria with extremely small genomes to study proteins essential for independent life
: *Mycoplasma genitalium* and *Mycoplasma pneumoniae*.
Principal Investigator: **Sung-Hou Kim, Lawrence Berkeley National Laboratory**

Center for Eukaryotic Structural Genomics
This Wisconsin-based center seeks to develop high-throughput methods for protein production, characterization and structure determination from *Arabidopsis thaliana*, a plant

The Joint Center for Structural Genomics
focus on novel structures from the roundworm *Caenorhabditis elegans* and on human proteins thought to be involved in cell signaling

The Midwest Center for Structural Genomics
unknown folds and on proteins from disease-causing organisms.

New York Structural Genomics Research Consortium
develop techniques to streamline every step of structural genomics.

Northeast Structural Genomics Consortium
target proteins from various model organisms--including the fruit fly, yeast, and the roundworm--and related human proteins.

The Southeast Collaboratory for Structural Genomics
two representative organisms--the roundworm *Caenorhabditis elegans* and its more primitive microbial ancestor, *Pyrococcus furiosus*.

Structural Genomics of Pathogenic Protozoa Consortium
This group aims to develop new ways to solve protein structures from organisms known as protozoans, many species of which cause deadly diseases

TB Structural Genomics Consortium
seeks to optimize the technical and managerial underpinnings of high-throughput structure determination and will develop a database of structures and functions

Microbial Genome Projects

Microbial Genome Projects

- Bacillus subtilis Genome Database (BSORF) Bioinformatics Ceter, Kyoto University and Nara Institute of Science and Technology
- Chlamydomonas Resource Center (Duke University, USA)
- Database of Genomes Analysed in NITE (DOGAN)
- Dictyostelium cDNA Database Dictyostelium discoideum cDNA Project (Dicty_cDB)
- *Dictyostelium Genome Sequencing Project*
- DOE Joint Genome Institute (JGI) Microbial Genomics, US
- E-coli genome project (K-12 and -157) (University of Wisconsin-Madison, US)
- GenoBase: Escherichia coli Genome Database Nara Institute of Science and Technology
- Genome Analysis Project Japan on E. coli (GenoBase)
- Genome Database for Cyanobacteria (CyanoBase) Kazusa DNA Research Institute
- Genome Information Broker (GIB) DNA Data Bank of Japan (84 microbes as of May 2002)
- Genome to Proteins and Functions
- GOLD: Genomes OnLine Database HomePage by Integrated Genomics Inc., US
- MagnaportheDB
- Malaria Full-Length cDNA Database (Plasmodium falciparum) (Institute of Medical Science, *The University of Tokyo, Japan*)
- Microbial Genome Database for Comparative Analysis (MBGD)
- PEDANT: Genome Analaysis and Annotation by MIPS, Germany
- Profiling of E.coli Chromosome (PEC)
- Saccharomyces Genome Information Server
- Synechocystis PCC6803 Gene Annotation Database (SYORF) Bioinformatics Ceter, Kyoto University and Cyanobacteria Research Community
- The Institute for Genomic Research

Microbial Genome Projects: Ex-- *Rhodopseudomonas palustris*

Strain: CGA009
Size: 5.47 Mb
Contigs: 2
Contig Sizes: 5.46 Mb (circular chromosome), 8.9 Kb (plasmid)
Reads: 117,398
Coverage: ~10.7X based on 500bp readlength
Currently: Completed 12/13/01

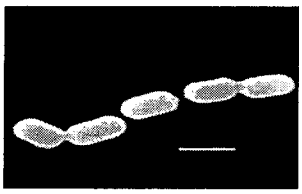
Rhodopseudomonas palustris

Who is *Rhodopseudomonas palustris?*

1. purple non-sulfur phototrophic bacterium
2. found in soils and water
3. converting sunlight to cellular energy
4. absorbing atmospheric carbon dioxide and converting it to biomass.
5. degrade and recycle a variety of aromatic compounds
6. convert nitrogen gas into ammonia,
7. degrading a variety of carbon-containing compounds (including sugars, lignin monomers, and methanol)

BRIC

Microbial Genome Projects: *Lactobacillus gasseri*



Lactobacillus gasseri

Strain: ATCC33323
Size: 1.8 Mb
Release: Version 1
Total Scaffold Size: 2.0 Mb

1. Lactobacilli are normal inhabitants of the gastrointestinal tract of man and animals
2. Beneficial roles including immunomodulation, interference with enteric pathogens, and maintenance of a healthy intestinal microflora
3. A variety of probiotic activities and roles
   ---reduction of fecal mutagenic enzymes
   ---adherence to intestinal tissues
   ---stimulation of macrophages
   ---production of bacteriocins
4. Expected to provide key insights to the survival, roles, and potential benefits of this group of commensal organisms
5. More amenable to DNA introduction and manipulation
   →useful in the functional genomic analysis of this species

BRIC

Microbial Genome Databases / portal sites:

TIGRhttp://www.tigr.org/tdb/mdb/mdbcomplete.html
GOLDhttp://216.190.101.28/GOLD/
DDBJhttp://gib.genes.nig.ac.jp/
EBIhttp://www.ebi.ac.uk/genomes/
MBGDhttp://mbgd.genome.ad.jp/
NCBIhttp://www.ncbi.nlm.nih.gov/PMGifs/Genomes/bact.html
PIRhttp://pir.georgetown.edu/pirwww/search/genome.html
STDdbhttp://www.stdgen.lanl.gov/
InfoBIOGENhttp://www.infobiogen.fr/doc/data/complete_genome.ht
MAGPIEhttp://www-fp.mcs.anl.gov/~gaasterland/genomes.htm
ICBShttp://www.cbs.dtu.dk/services/GenomeAtlas/
KEGGhttp://www.genome.ad.jp/kegg/catalog/org_list.html
NIADhttp://www.niaid.nih.gov/dmid/genomes/genome.htm
Completely Sequenced Genomeshttp://linkage.rockefeller.edu/wli/seq/
CyanoBasehttp://www.kazusa.or.jp/
ParasiteDBhttp://www.ebi.ac.uk/parasites/paratable.html

NCBI
EBI
DDBJ
CBS
Los Alamos
INFOBIOGEN
MBGD MICROBIAL GENOME DATABASE
parasite-genome

BRIC

Microbial Cell Research

1. Biological Basics Goal: Determine and characterize the minimum set of genes and corresponding gene products necessary to sustain a simple free-living microbial cell, express the genes to produce the relevant proteins, and determine their structure.

2. Functional Foundations
Goal: Determine the physiological and biochemi-cal functions of the gene and specific bioprocesses using standard biochemical techniques and structural/ computational biology.

3. Modeling Interactions
Goal: Use high-end computing to model gene-gene, gene-protein, and protein-protein interac-tions as well as the internal biochemistry of the cell.

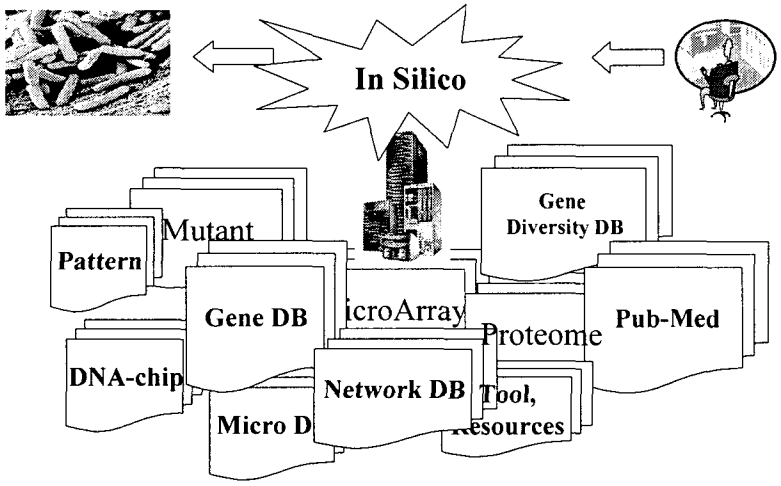
4. Regulation and Manipulation
Goal: Use gene-protein manipulation to enhance or suppress various cell functions.

5. Functional Expression
Goal: Focus on functions that are relevant to DOE goals (e.g., bioremediation, carbon sequestration, and sustainable energy production).

BRIC

In Silico Microorganism

통합 검색 분석, 유전자 기능 예측 및 **simulation**, 원하는 미생물 디자인



BRIC

In Silico Microorganism

Genomatica To develop " Next-Generation" *In Silico* Technology

Start: Jan 3, 2002
Organism: Microbe
Fund: Department of Energy—$ 2 Million
Mission: Generating of integrated models of metabolism and metabolic regulation

"Cybernetic Cells" explores the movement by biologist and companies to simulate the biochemistry of living cell, with the ultimate goal of finding a way to perform virtual experiment In silico that can speed up the discovery of new drugs and Reduce their cost

BRIC

Microbial genome sequencing: applications and future directions

IMMEDIATE

DNA
- IDENTIFY GENES, ORFS, ETC.
- TRANSCRIPTION/TRANSLATION START/STOPS
- NUCLEOTIDE COMPOSITION
- "WORDS", CONSENSUS MOTIFS
- CODON USAGE
- GENOME STRUCTURE

GENES
- ORTHO-/PARA-/HOMOLOGS
- GENE FAMILIES
- PREDICTED CELL LOCALIZATION
- MOTIFS

POST-GENOMIC STUDIES

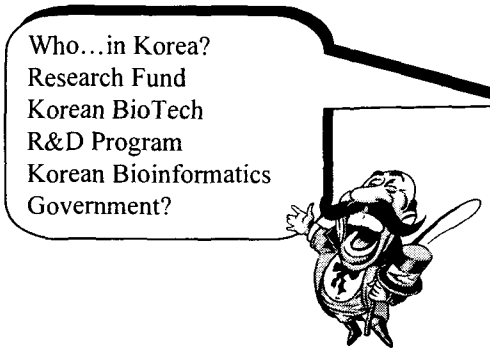
LAB EXPERIMENTS
- GLOBAL GENE EXPRESSION
- GLOBAL PROTEIN ANALYSIS
- GLOBAL MUTAGENESIS (knock-outs, reporter genes)
- PROTEIN OVEREXPRESSION
- SCREENS/ASSAYS (gene function, drug design, virulence)

"IN SILICO STUDIES"
- COMPARATIVE GENOMICS
- 3-D STRUCTURAL PREDICTIONS
- BIOCHEMICAL PATHWAYS
- GENE EXPRESSION HIERARCHY

FUTURE
Comprehensive understanding of:
- PHYSIOLOGY
- GENETICS
- EVOLUTION
- BIOCHEMISTRY

BRIC

Trends in Microbial Genomics and Bioinformatics



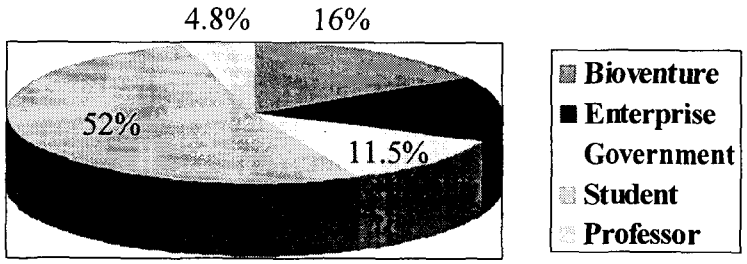
1. Introduction
2. Microbial Genome
3. Korean Bioinformatics
4. Bioinformatics Market
5. Education
6. BRIC's Bioinformatics

BRIC

Korean Bioinformatics

Who.....in Korea?

Average number of participants in the workshops held in 2001 by Korean Society for Bioinformatics

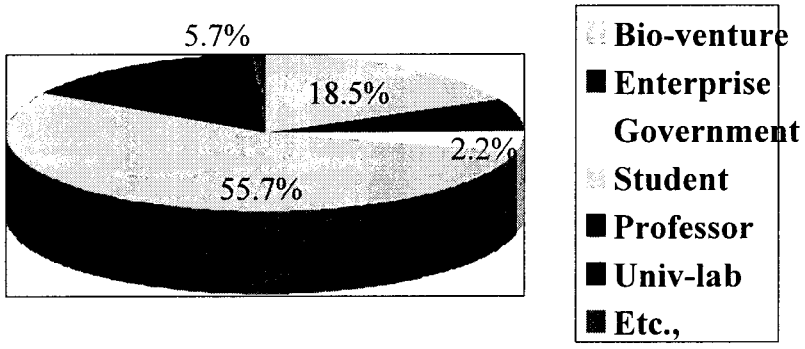


Provided by KSBI

Korean Bioinformatics

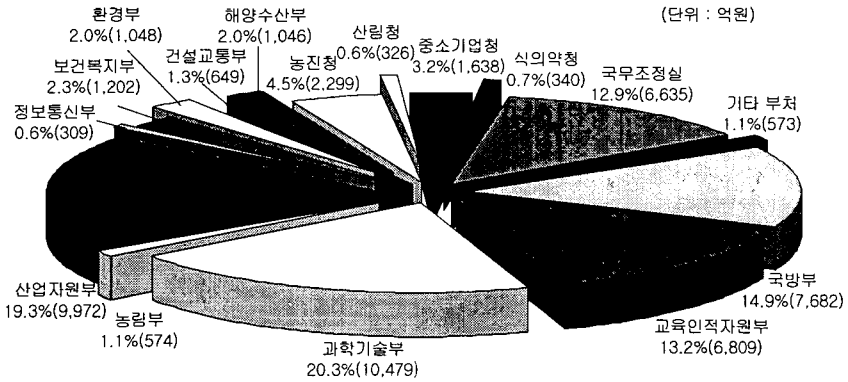
Who.....in Korea?

2002 The First workshop : Participants



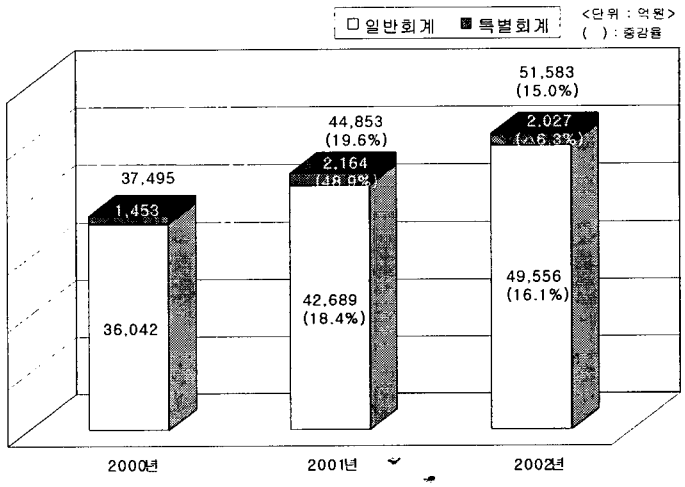
Provided by KSBI

2002년도 부처별 연구개발예산 분포



2002년 R&D 예산 : 5조 1,583억원

정부연구개발예산의 추이(2000~2002년)



자료: 과학기술부 한국과학기술기획평가원

미래 유망 신기술분야 예산편성 현황

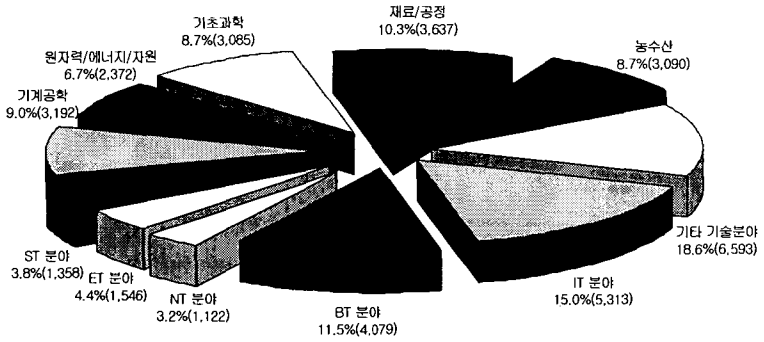
(단위:억원)

구분	2000년	2001년(A)	2002년(B)	증감 (B-A)	%
정보기술 (IT)	4,085	4,536	5,313	777	17.1
생명공학기술(BT)	2,462	3,353	4,079	726	21.7
나노기술 (NT)	300	425	1,122	697	163.7
환경기술(ET)	955	1,328	1,546	218	16.4
우주항공기술(ST)	721	1,212	1,358	146	12.0
합계	8,563	10,854	13,418	2,564	23.6

자료: 과학기술부 한국과학기술기획평가원

2002년도 기술분야별 예산분포

(단위 : 억원)



기술분류 가능 2002년 R&D 예산 : 3조 5,387억원

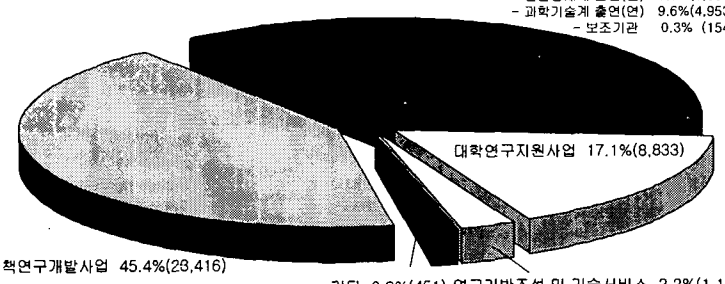
자료: 과학기술부 한국과학기술기획평가원

2002년도 기능별 연구개발예산 분포

(단위 : 억원)

연구기관운영지원 34.4%(17,749)

- 국립 연구 및 시험기관 7.7%(3,959)
- 부처 소속 출연(연) 13.8%(7,125)
- 인문경제계 출연(연) 3.0%(1,558)
- 과학기술계 출연(연) 9.6%(4,953)
- 보조기관 0.3% (154)

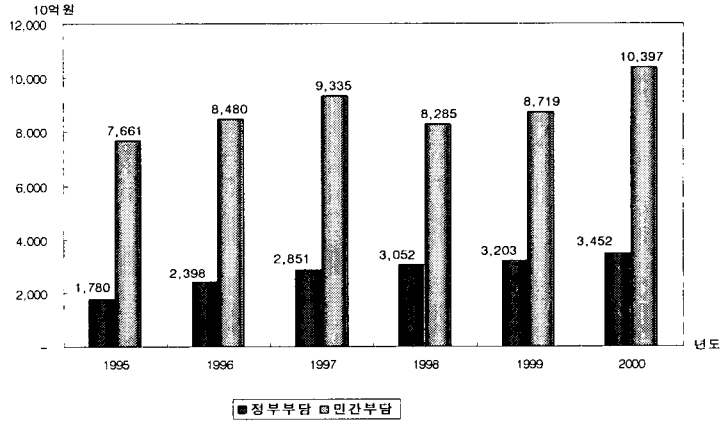


- 특정연구개발 11.5%(5,954)
- 산업기술개발 19.1%(9,832)
- 원자력/에너지 1.7% (879)
- 농림/해양 1.8% (951)
- 환경 1.4% (720)
- 건설/교통 1.5% (789)
- 보건/의료 2.3%(1,203)
- 국방 5.3%(2,724)
- 기타 0.7% (365)

2002년 R&D 예산 : 5조 1,583억원

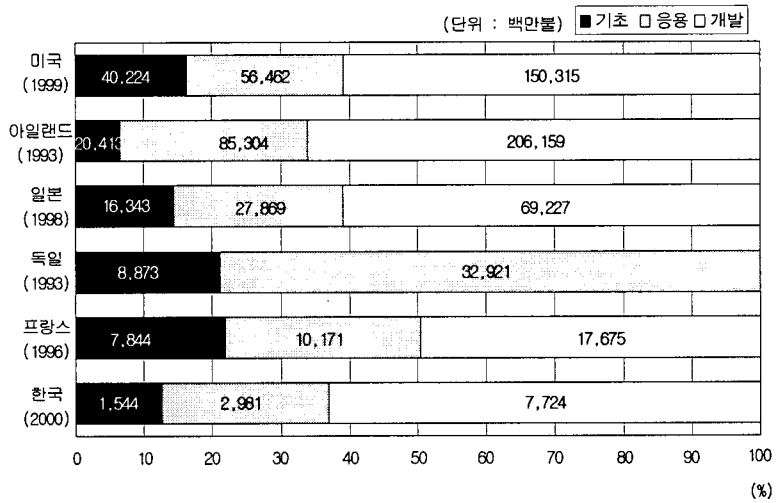
자료: 과학기술부 한국과학기술기획평가원

총 연구개발비의 재원별 추이(1995~2000년)



자료: 과학기술부 한국과학기술기획평가원, 「2001 과학기술연구활동조사보고」, 2001.

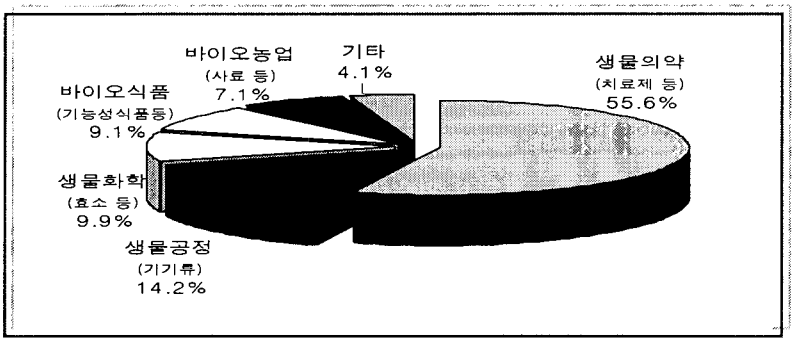
주요국의 성격별 총 연구개발비 구성 국제비교



<자료원>: 과학기술부 한국과학기술기획평가원, 「2001 과학기술연구활동조사보고」, 2001.

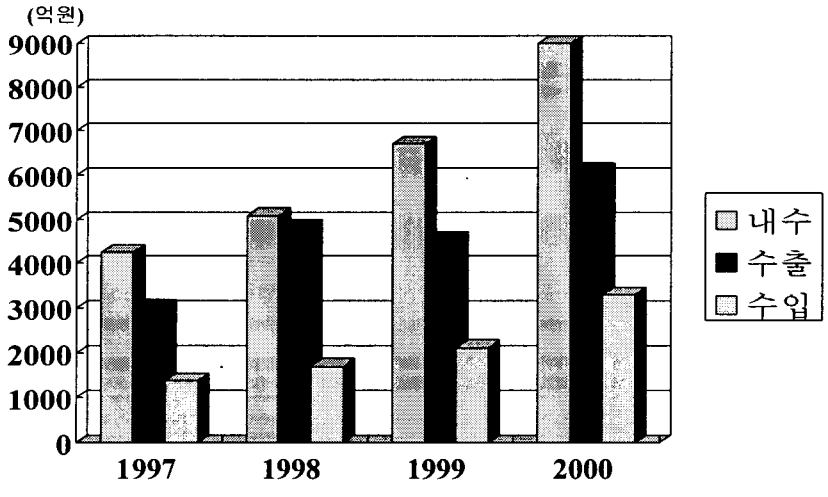
2000년 국내 바이오산업의 분야별 구성비

분야	구성비	분야	구성비
생물의약(치료제등)	55.6%	바이오식품	9.1%
생물공정(기기류)	14.2%	바이오 농업(사료)	7.1%
생물화학(효소등)	9.9%	기타	4.1%



자료: 생물산업협회, 국내 생물산업 실태조사표

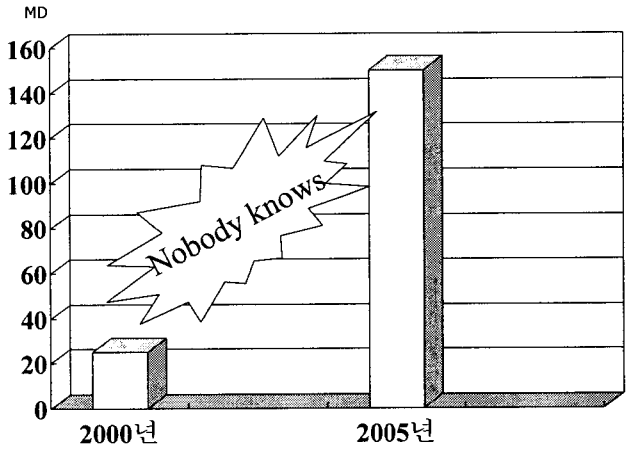
국내 바이오 산업 시장규모



자료: 생물산업협회, 국내 생물산업 실태조사표

BRIC

Korean Bioinformatics Market(2000~2005)



From LG Economic Research Institute

BRIC

바이오 산업 정부투자 규모

년도	1994	1996	1998	1999	2000	2001	2002
억원	536	1234	1115	1608	2462	3238	4500

▶ 2000년을 기준으로 대폭 증가

▶ **R&D** 예산대비 바이오분야 비중은 **8%**수준

- 94년(3.5%)→98년 (5%)→2001년(8%)
- 98년 이후 증가율: 바이오분야 46%, 정부전체 :19.4%

▶ **2002**년 기준 정부총액 투자규모:

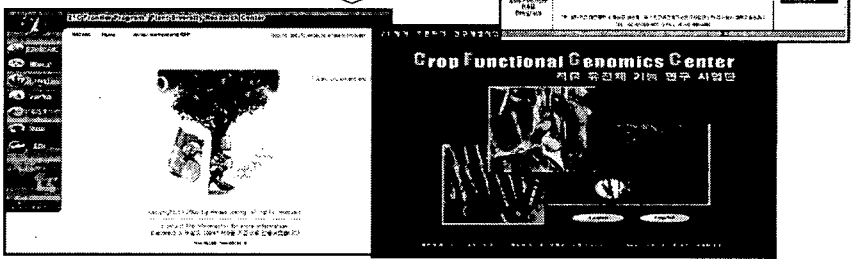
- Amgen(8.5억$/년)의 1/3수준

Bioinformatics R&D Pogram

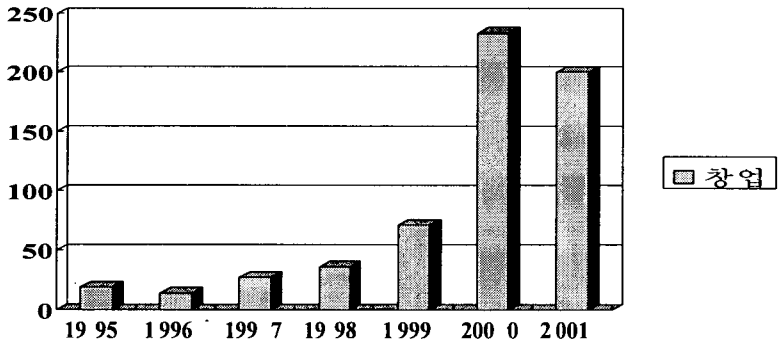
Genomics and Proteomics R&D Program

: Government grants of US$10 million per year for 9 years

>> The Center for functional Analysis of Human Genome
>> Crop Functional Genomics Center
>> Plant Diversity Research Center
>> Human Proteomics Center
>> Drug Discovery Center
>> Medical Genomics program



Korean Bio-Venture



•현재 약 600여개로 추정
•수익모델 악화로 수적양적인 구조조정이 이루어지고 있음

BRIC

Korean Bio-Venture

Bioventure	*Research fields*
Macrogen :	DNA chip, genome analysis, Bioinformatics solution
Neodin :	Medical diagonosis kit
DigitalGenomics :	Medical diagonosis kit
Bioinfomatxs:	Genome and proteome analysis tool,
VectorKorea:	Gene therapy
IDR:	Protein prediction and modeling, New drug development
DNAlink:	New drug development
Bioneer:	Bio-vendor, sequencing machine and new drug development
Genomictree :	DNA chip and sequencing service
Toolgen:	New drug development
Inbionet:	Genome analysis and data mining,
Proteogen:	Protein chip, Protein analysis machine
Badasoft:	Bioinformatics solution, System development
Cristalgenomics:	New drug development, structural Proteomics

Etc...,

From Korean Bioventure Association

BRIC

Korean Bioinformatics

Industrial Research Fields-Enterprise

- **Samsung SDS:** Genome analysis portal solution, Network solution, Analysis tools of Proteome and DNA chip
- **LG Chemicals:** New Drug discovery, Gene discovery Genome research.
- **SK:** New Drug discovery and Gene therapy
- **Hanhwa:** New Drug discovery in cardinal and nerves diseases

SWOT Analysis

Strength(S)
Information Technology Infra
Need Power
Man-Power in Biotechnology

Opportunity(O)
Large scale market
Start step
Many kinds of research fields

Weakness(W)
A few specialist in bioinformatics
Government Support
A few special Institute
Infra in Biotechnology

Treat(T)
Competition with an advanced Country
Endeavor of multinational Enterprise

What does Government do?

- Providing scientific, technical, organizational and policy advice
- Supplying information and information products
- Promoting new technologies, strategies and tools
- Providing training and/or information on training opportunities
- Bringing people and institutions together at national and sub-regional levels
- Assisting in writing project proposals and obtaining funds
- Implementing, coordinating and publishing research
- Allocating direct financial support to infrastructure development and activities

BRIC

Trends in Microbial Genomics and Bioinformatics

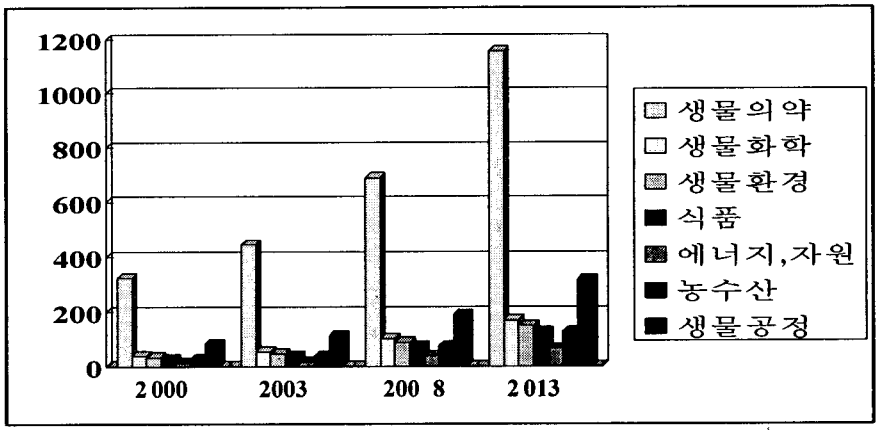
1. Introduction
2. Microbial Genome
3. Korean Bioinformatics
4. Bioinformatics Market
5. Education
6. BRIC's Bioinformatics

Market
Changing Marketplace
Pharma and Biotech
Drug discovery
Prospects

BRIC

세계 바이오산업 시장규모

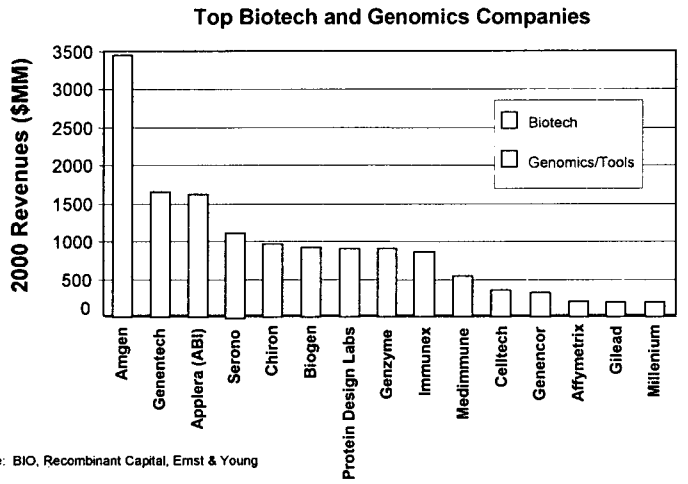
(단위: 억불)



* 자료 : OECD, *Biotechnology & Trade*, 1997
CRC, *U.S. Biotechnology Product Sales Forecast*, 1997

Biotech and Genomics Companies(2000 Revenue)

The Biotech landscape is dominated by a few key players, a dozen smaller companies, and hundreds of small start-ups.



Total Industry Revenue, 2000 = $ 24.9 B

Total Number of Biotech Companies, 2000 = 1273

Number of *Profitable* Biotech Companies, 2000 = 35

Source: BIO, Recombinant Capital, Ernst & Young

The Changing Marketplace of Bioinformatics

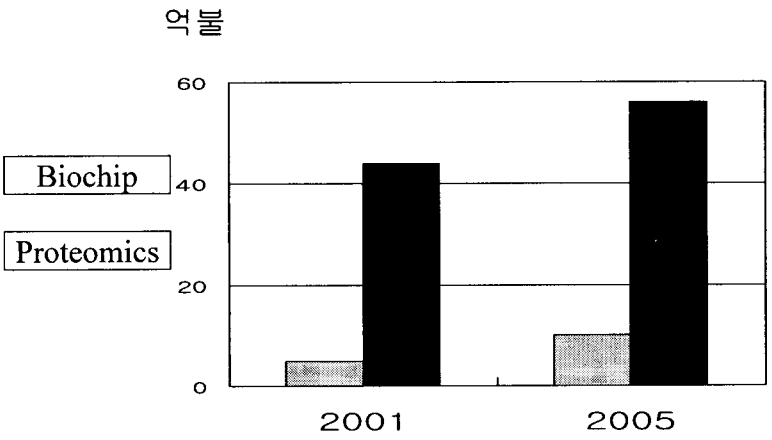
Several roles are evolving for bioinformatics in life science

- Automation of biological and chemical research for the drive toward high-throughtput procedures
- Providing tools for the capture, management and analysis of disparate data
- Facilitating the exchange and dissemination of information between isolated groups in large organizations

Emerging niches

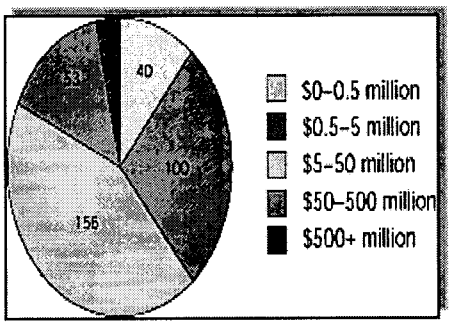
- **Poviding software packages for data visualization, interpretation, and Analysis**
  : ex—Spotfire(Cambridge, MA), MDL(SanLeandro, CA), Silicon Genetics(SanCarlos, CA)
- **Offering comprehensive bioinformatics platforms accessible to life scientists over internet**
  : ex—Doubletwist(Oakland, CA), Compugen/LabonWeb(Tel Aviv, Israel), e-Bioinformatics
- **To rent out Data architectures/infrastructure and bioinformatics expertise for data integration**
  : ex—LionBioScience(Heidelberg, Germany), IBM, Motorola
- **Biological Research-based Companies are offering tools along with their contents**
  :ex—Incyte Genomics, Celera Genomics, Curagen(New Haven, CT)

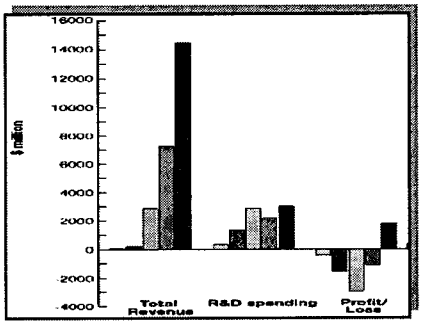
Huge Market Opportunities



Citation by **Genomic Solutions, 2001**

Biotechnology Companies by Revenue





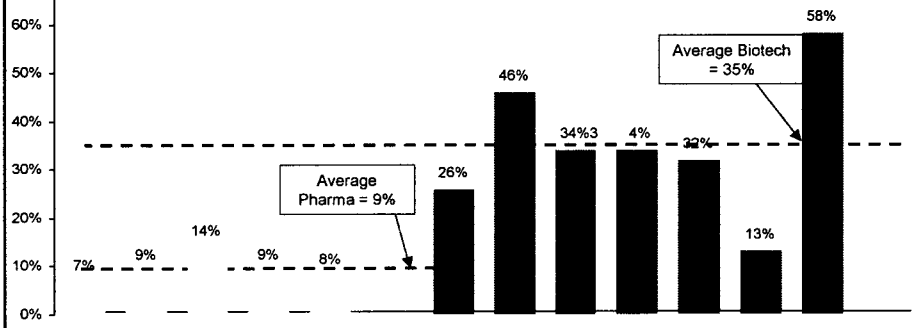
Distribution of publicly traded biotechnology companies according to revenue generated.
Total revenue, R&D spending, and profit/loss for the five categories of company described

Nature Biotechnology 19, 407 – 412 (2001)

Pharma and Biotech in R&D

Biotech's high % of R&D expenditures reflect a focus on innovation and a tendency to license out products as they near the commercialization phase.

R&D Expenditures as a Percent of Sales, 1999



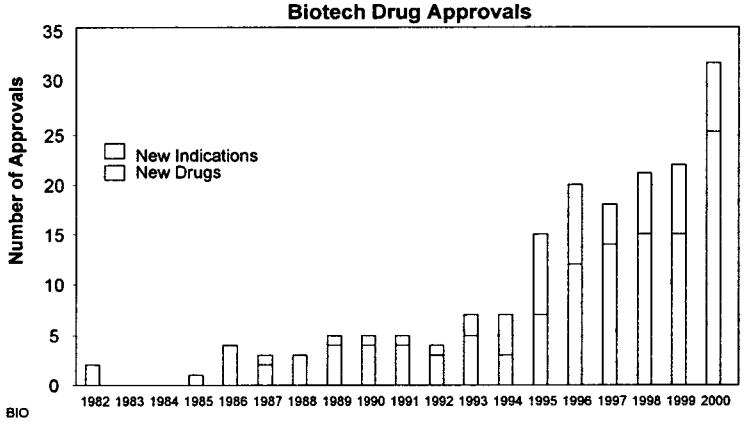
	Merck	Pfizer	GSK	J & J	BMS	Amgen	Genentech	Biogen	Alza	Chiron	Genzyme	Immunex
R&D Expenditures per Employee ($000):	$31	$39	$36	$24	$26	$119	$145	183	$102	$58	$21	$123

Source: Annual reports, Ernst & Young

Change of Drug Market

The number of Biotech drugs on the market has skyrocketed since 1994, and the trend is expected to continue.

Biotech Drug Approvals



Source: BIO

An additional 350 biotech medicines targeting over 200 diseases were in late-stage development in 2001, promising to continue the upward trend in approvals.

Mergers and acquisitions *between Pharma and Biotech*

The distinction between Pharma and Biotech is becoming blurred as the industry consolidates.

Many of the top Biotechs are partially or entirely owned by traditional Pharma companies

- **Genentech and Roche - Chiron and Novartis**
- **Immunex and AHP - Centocor and J&J**

Mergers and acquisitions among Biotechs are also common

- **Millennium Pharmaceuticals and LeukoSite.**
- **Genzyme and Cell Genesys, also Peptimmune.**
- **MedImmune and U.S. Bioscience.**
- **ALZA and Sequus Pharmaceuticals .**

Few very small integrated pharmaceutical companies exist, as economies of scale in drug development, sales, and marketing are required to drive down costs and maintain growth rates

Biotechnology company casualties during 2000

Table 3. Biotechnology company casualties during 2000

Mergers
Algos Pharmaceutical and Endo Pharmaceuticals
Alfalix Biopharmaceuticals and NPS Pharmaceuticals (1999)
Aquila Biopharmaceuticals and Antigenics
Biomatrix with Genzyme Tissue Repair and Genzyme Surgical Products, to form Genzyme Biosurgery
GelTex Pharmaceuticals and Genzyme
Biopool and Xtrana
Celtech Chiroscience and Medeva, forming Celtech Group
Coulter Pharmaceutical and Corixa
Creative Biomolecules with Ortogeny and Reprogenesis, to form Curis
Epitope and STC Technologies
Evotec BioSystems and Oxford Asymmetry
Life Technologies and Invitrogen
Lunar and GE Medical Systems

Acquisitions
Accumed, by Ampersand
Agritope, by Exelixis
Anesta, by Cephalon
Cambridge NeuroScience, by CeNeS
Cetrix Pharmaceuticals, by Insmed
ChiRex, by Rhodia
CliniChem, by Biochem
Crescendo Pharmaceuticals, by Alza
Trega Biosciences, by Lion Bioscience
Spiros, by Dura Pharmaceuticals
Oxford Molecular, by Pharmacopeia
Pathogenesis, by Chiron
LJL Biosystems, by Molecular Devices
North American Vaccine, by Baxter
Quadrant Healthcare, by Elan
Liposome, by Elan
Dura Pharmaceuticals, by Elan

Name changes
CytoTherapeutics became StemCells
Energy Biosystems became Enchira Biotechnology
Ethical Holdings became Amarin
Meridian Diagnostics became Meridian Bioscience
NovoPharm Biotech became Viventia Biotech
Peptide Therapeutics became Acambis
QLT PhotoTherapeutics became QLT
UroGen changed name to GenStar (and relisted)
Vanguard Medica changed name to Vornalis
Medi-Ject changed name to Antaros Pharma

Delisted
Amarillo Biosciences
BioSante Pharmaceuticals
Cadus Pharmaceutical
Clinicor
Ecogen
GalaGen
Hybridon
Scotia Holdings
Unigene

Relistings
Antex Biologics
Epoch Pharmaceuticals

Nature Biotechnology 19, 407 - 412 (2001)

BRIC

Mergers and acquisitions *between Pharma and Biotech*

Alliances and licensing have been instrumental in Biotech industry growth, but are becoming more central to Pharma's business model as well.

Advantages for Pharma

- Efficient innovation
- Access to new technologies
- Diversification of development risk
- Enhanced product pipelines
- Downside: lack of control over pipeline

Pharma needs Biotech more now than in the past for drug discovery and new products as their scale makes sustainable growth difficult

Advantages for Biotechs

Access to resources

Marketing expertise

Established sales force

Downside: revenue-sharing and postponement of value-chain integration

Biotechs need pharma less due to multiple sources for manufacturing and marketing, and downstream integration.

BRIC

Biotechnology Financing, 2000~2001

Biotechnology financing, 2000–2001

Quarter	Total (billions of US $)	US	Europe*
Q3 2001	2.46	1.27	1.06
Q2 2001	5.39	4.54	0.59
Q1 2001	2.34	1.91	0.24
Q4 2000	10.23	8.69	1.53
Q3 2000	8.78	6.12	2.52
Q2 2000	4.65	3.17	1.30
Q1 2000	13.09	11.71	0.52

*For all offerings in Austria, Belgium, Denmark, Finland, France, Germany, Hungary, Iceland, Ireland, Israel, Italy, Netherlands, Norway, Spain, Sweden, Switzerland, and UK
Source: BioCentury.

BRIC

Next Waves of Innovation in Drug Discovery

Table 1. Three waves of innovation in drug discovery

Innovation	Tools/technologies	Representative companies
First wave		
Genomics (target discovery, antisense therapeutics, and ultimately gene therapy)	Gene / genome sequencing and expressed sequence databases	Incyte (Palo Alto, CA); Celera (Rockville, MD); Gemset (Paris); Genome Therapeutics (Waltham, MA)
	Full-length cDNAs/expression	Clontech (Palo Alto, CA); Stratagene (La Jolla, CA); Life Technologies (Rockville, MD)
	Functional genomics	Curagen (New Haven, CT); Millennium (Cambridge, MA) and Pharmagene (Cambridge, UK)
	RNA expression profiling	Affymetrix (Santa Clara, CA); Gene Logic (Gaithersburg, MD); Rosetta Inpharmatics (Kirkland, WA)
	Transgenesis / directed mutagenesis	Lexicon Genetics (Woodlands, TX); Genome Systems (St. Louis, MO)
	Antisense Tetanology	Isis Pharmaceuticals (Carlsbad, CA); Hybridon (Cambridge, MA)
	SNP databases	Variagenics (Cambridge, MA); Genaissance (New Haven, CT); Celera
	Genetic mapping / disease genes	Oxagen (Boulder, CO); Myriad (Salt Lake City, UT); Decode (Reykjavik, Iceland)
	Genetic diagnostics	Quest (Tarrytown, NJ); Roche (Nutley, NJ); Diadexus (Palo Alto, CA); Affymetrix
Second wave		
Proteomics (target validation, drug screening, antibody therapeutics, and protein therapeutics)	Protein databases and protein expression analyses	Oxford Glycosciences (Oxford, UK); Large Scale Biology (Vacaville, CA); Proteome (Cambridge, MA)
	Protein expression technologies and protein therapeutics	Amgen (Thousand Oaks, CA); Genentech (S. San Francisco, CA); Human Genome Sciences (Rockville, MD); Chiron (Emeryville, CA); Genetics Institute (Cambridge, MA); Lonza (Slough, UK)
	Directed evolution	Maxygen (Redwood City, CA); Phylos (Lexington, MA);
	Antibody engineering	Celltech/Medarex (Leatherhead, UK); Abgenix (Fremont, CA); Cambridge Antibody Technology (Cambridge, UK)
	High-throughput screening	Aurora Biosciences (La Jolla, CA); Cambridge Drug Discovery (Cambridge, UK)
	Protein interaction databases	Proteome, MDS-Proteomics (Etobicoke, QC, Canada); Celera
	Affinity selection	Neogenesis (Cambridge, MA); MDS-Proteomics
	Protein pathways / protein chips	Zyomyx (Hayward, CA); Combimatrix (Seattle, WA); Cytogen (Palo Alto, CA); Sense Proteomics (Cambridge, UK)
Third wave		
Molecular design (protein therapeutics, antibodies, and small molecules)	Protein structure determination	Structural Genomix (San Diego, CA); Syrrx (San Diego, CA); Astex (Cambridge, UK)
	Protein homology modeling	Structural Bioinformatics (San Diego, CA); Geneformatix (San Diego, CA)
	Protein engineering	Surrozis (Redwood City, CA); Sangamo (Richmond, CA); Cambridge Antibody Technology (Cambridge, UK)
	Structure-based small molecule design	Vertex Pharmaceuticals (Cambridge, MA); De Novo (Cambridge, UK)
	Molecular design tools	Molecular Simulations (San Diego, CA); Tripos (St. Louis, MO)

Nature Biotechnology 19, 207 - 209 (2001)

BRIC

Prospects

New trends in Biology

- **New research trend in biology.**
- **Systems biology**
  - A unique approach to the study of genes and proteins which has only recently been made possible by rapid advances in computer science.
  - Unlike traditional science which examines single genes or proteins, systems biology studies the complex interaction of all levels of biological information:
    - Genomics DNA, mRNA, proteins, functional proteins, informational pathways and informational networks to understand how they work together.

BRIC

Prospects

New needs in Biology

➢ Changes of Meaning

–The role of Bioinformatics will be critical and expand.

- In the narrow view: tools for storing, processing, and searching data.
- In the broad view: discipline for creating new knowledge in Biology.

➢ Into a more quantitative science.

- sophisticated mathematical, statistical, and data modeling approach.

➢ More integrative Approach

- Mathematician, Statistician, Physicist, and Computer Scientist
- Engineering, Linguistics, Fluidics, Robotics, ...
- Linguistics:
  -Develop languages for Biology in the view of formalism.

BRIC

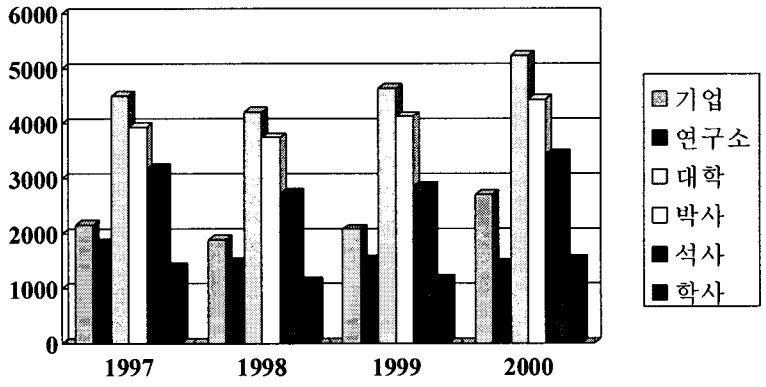
Trends in Microbial Genomics and Bioinformatics

1. Introduction
2. Microbial Genome
3. Korean Bioinformatics
4. Bioinformatics Market
5. Education
6. BRIC's Bioinformatics

보도자료

- 연합뉴스(01/5/29): 바이오산업 핵심 생물정보학 인력 "절대부족"
- 매일경제(01/6/4): 생물정보학 교육인력 부족
- 연합뉴스(01/07/19): 산자부, 서울대 BT인력 재교육 나선다
- 전자신문(01/8/20): 바이오업계 인력난 몸살
- 과학신문(01/8/29): 바이오인포매틱스 인력양성 차질
- 한국경제(01/10/27): 일부 바이오업체, 사업다각화 고급인력 확충나서
- 약업신문(01/11/15): 산자부, BT 인력 육성에 5,000억원 투입

전문인력



- ▶ 대학: 5224명, 기업:2,701명, 연구소: 1,467명
- ▶ Genomics, Proteomics, Bioinformatics분야:
- ▶ 2005년 인력 수급차질 예상: 수요—9470명, 공급—3080명(6,300명 정도 부족예상)

자료: 2002년 생명공학육성시행 계획

국내 생물정보학 교육과정

학교	구분	소속	학사	석사	박사	개강년도	기타
부산대학교	협동과정	대학원		◆	◆	2000.5	국내최초
숭실대학교	학과설립	자연대	◆			2001.3	학부최초
서울대학교	협동과정	생명과학부		◆	◆	2002.3	다수의 교수진
KAIST	학과설립	공학부	◆	◆	◆	2002.9	학부, 대학원
중부대학교	학과설립	생명과학부	◆			2002.3	
국립보건원	단기과정					2001.6	1 회 실시 종료

향후 교육과정

1. 대학원 협동과정 중심으로
2. 특화된 분야 과정
3. 재교육—수강자에 따라 다양한 프로그램개발
4. **On-line** 교육 활성화
5. 전문강사 확보를 위한 장기프로그램

대표적인 국외 생물정보학 교육과정

Some degree programs in bioinformatics/computational biology

Table 1. Some degree programs in bioinformatics/computational biology

School	Degree	URL
Boston University (Boston, MA)	MS/PhD	http://bioinformatics.bu.edu/
George Mason University (Fairfax, VA)	MS/PhD	http://www.science.gmu.edu/~michaels/Bioinformatics
Georgia Tech (Atlanta, GA)	MS	http://www.biology.gatech.edu/bioinformatics/
Iowa State University (Ames, IA)	PhD	http://www.bcb.iastate.edu/
Medical College of Wisconsin/Marquette Univ. (Milwaukee, WI)	MS	http://goliath.ahc.mcw.edu/AP/announcementsinitial.html
Rensselaer Polytechnic Institute (Troy, NY)	BS	http://www.rpi.edu/dept/bio/mlo/bioinformatics.html
Rutgers University (New Brunswick, NJ)	PhD	http://cmb.rutgers.edu/
University of the Sciences in Philadelphia (Philadelphia, PA)	MS	http://www.usip.edu/bioinformatics/

Degree programs that offer specialization/emphasis in bioinformatics

Table 2. Degree programs that offer specialization/emphasis in bioinformatics

School	Degree	URL
Baylor University (Waco, TX)	BS/Informatics	http://ecswww.baylor.edu/ecs/computer_science/csundergrad.htm
North Carolina State Univ. (Raleigh, NC)	MS, PhD/Genomics	http://genomics.ncsu.edu/
University of California, San Diego	PhD/Comp. Sci., Eng., Math., Chemistry, Physics, etc.	http://www.ogsr.ucsd.edu/bioinformatics/program.htm
University of Pennsylvania (Philadelphia, PA)	BS/MS/PhD, Computer Sci., Biotech., Biology, Math	http://www.cbil.upenn.edu/UPCB/
University of Pittsburgh/ Carnegie Mellon (Pittsburgh, PA)	PhD/Computer Sci., Biol. Sci., Chemistry, Physics, etc.	http://www.cs.pitt.edu/keck/program.html
Washington University (St. Louis, MO)	PhD/Medicine, D.Sc./Biomed. Eng.	http://www.ibc.wustl.edu/CMB/

Nature Biotechnology 19, 285 - 286 (2001)

BRIC

Trends in Microbial Genomics and Bioinformatics

Bioinformatics Team
대외 활동사항
KSDB
Service Items

1. Introduction
2. Microbial Genome
3. Korean Bioinformatics
4. Bioinformatics Market
5. Future of Bioinformatics
6. BRIC's Bioinformatics

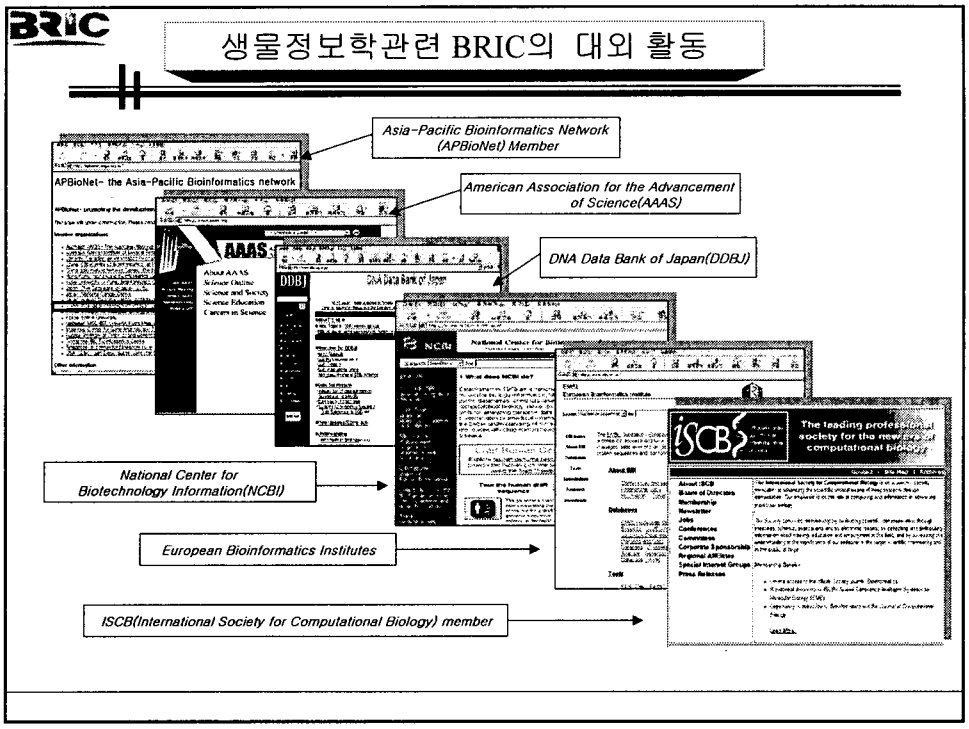
BRIC

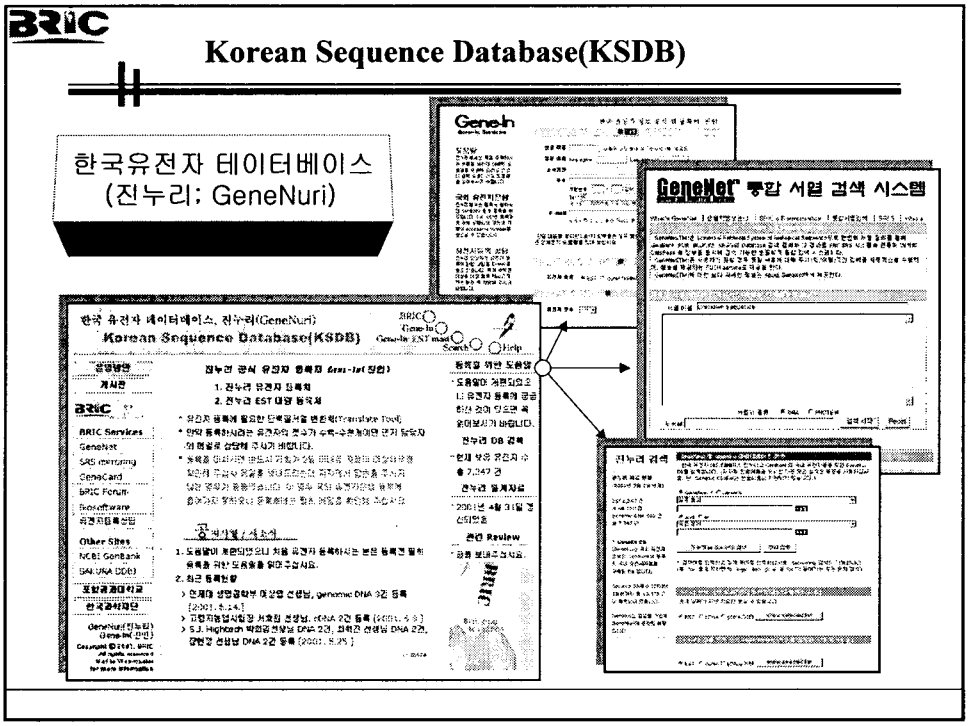
BRIC's Bioinformatics Team

Biology (생물학팀)
- -Molecular Biology
- -Genetics
- -Microbiology
- -Proteomics
- -Functional Genomics

Statistics (통계팀)
- -Statistics & Mathematics
- -Applied Statistical Modeling
- -Biological Data Classification
- -Clustering

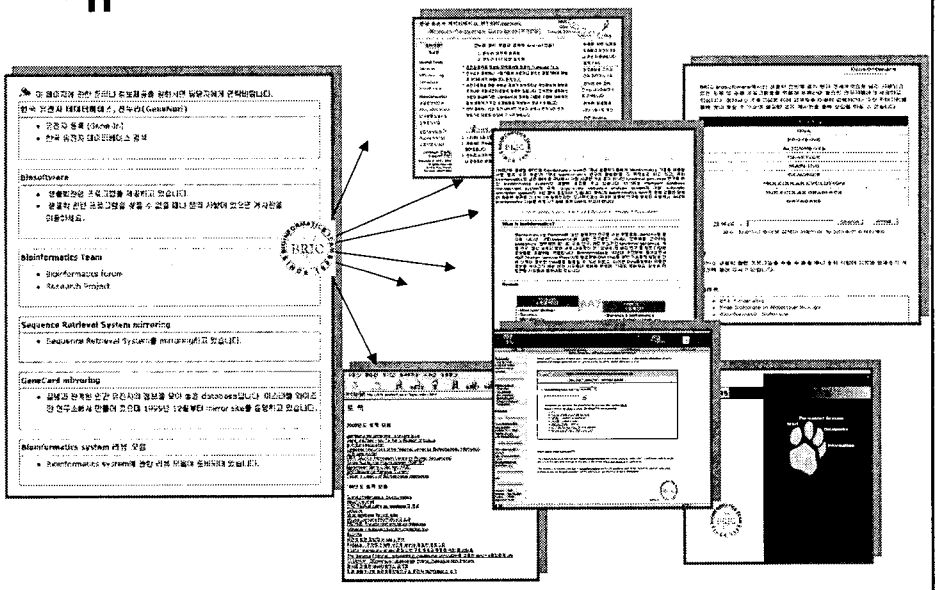
Computation (전산팀)
- -Bioperl
- -BioJAVA
- -Biopython
- -Biocorba
- -AI (Artificial Intelligence)
- -Algorithm
- -Natural Language Processing


BRIC
생물정보학관련 BRIC의 대외 활동
Asia-Pacific Bioinformatics Network (APBioNet) Member
American Association for the Advancement of Science(AAAS)
DNA Data Bank of Japan(DDBJ)
National Center for Biotechnology Information(NCBI)
European Bioinformatics Institutes
ISCB(International Society for Computational Biology) member


BRIC
Korean Sequence Database(KSDB)
한국유전자 데이터베이스
(진누리; GeneNuri)
GeneNet 통합 서열 검색 시스템

BRIC

Service Items for Bioinformatics in BRIC



BRIC

과학기술부산하 한국과학재단지정 생물학전문연구정보센터

Biological Research Information Center