

인 용 호 박사

바이오인포매틱스(주), 연구소장

Tel. +82-2-3445-4533, Fax. +82-2-3445-1281

E-MAIL : inyh66@bioinfomatix.com

Address : 서울시 강남구 논현동 81-7 <우:135-818>

◆ 연구관심분야 :

Bioinformatics, Systems Biology

◆ 학 력 :

1989년 2월 이학사, 한양대학교 생화학과

1991년 2월 이학석사, 한양대학교 대학원 생화학과

2000년 2월 이학박사, 한양대학교 대학원 생화학 및 분자생물학과

◆ 주요경력 :

1989. ~ 1990. 한양대학교 생화학과 조교

1990. ~ 1997. (주)미원 중앙연구소 선임연구원

1997. ~ 1999. 미국 켄터키대 의대 생리학과 초빙연구원(Visiting Fellow)

2000. ~ 2000. 포항공대 생물학전문연구정보센터 총괄팀장

2000. ~ 현 재 바이오인포매틱스(주) 부설연구소 소장

2001. ~ 2001. 한양대학교 생화학 및 분자생물학과 강사

2001. ~ 현 재 한국생물정보학회 국제간사

2002. ~ 현 재 한양대학교 생화학 및 분자생물학과 겸임교수

◆ 학회 활동 / 수상 / Honors

2000 - 현재 한국생물정보학회 회원

2000 - 현재 JSBi(Japanese society of bioinformatics) member

2000 - 현재 ISCB(International society for computational biology) member

◆ 연구 실적 요약

• 학술잡지 논문발표 : 3 편

• 학술컨퍼런스 논문발표 : 9 편

• 초청강연, 초청세미나 : 다수

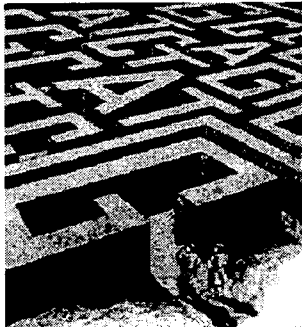
• 특허 : 5 건



한국생물정보학회
Korean
Society for
Bioinformatics

WELCOME TO THE KOREAN SOCIETY FOR BIOINFORMATICS

Genome Sequencing and Analysis of *Mannheimia succiniciproducens* MBEL55E

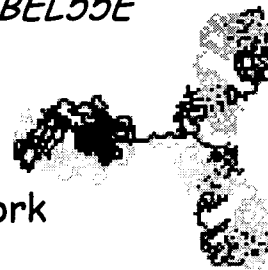


Bioinfomatix Inc.
R & D Institute
Yong-ho In
inyh66@bioinfomatix.com



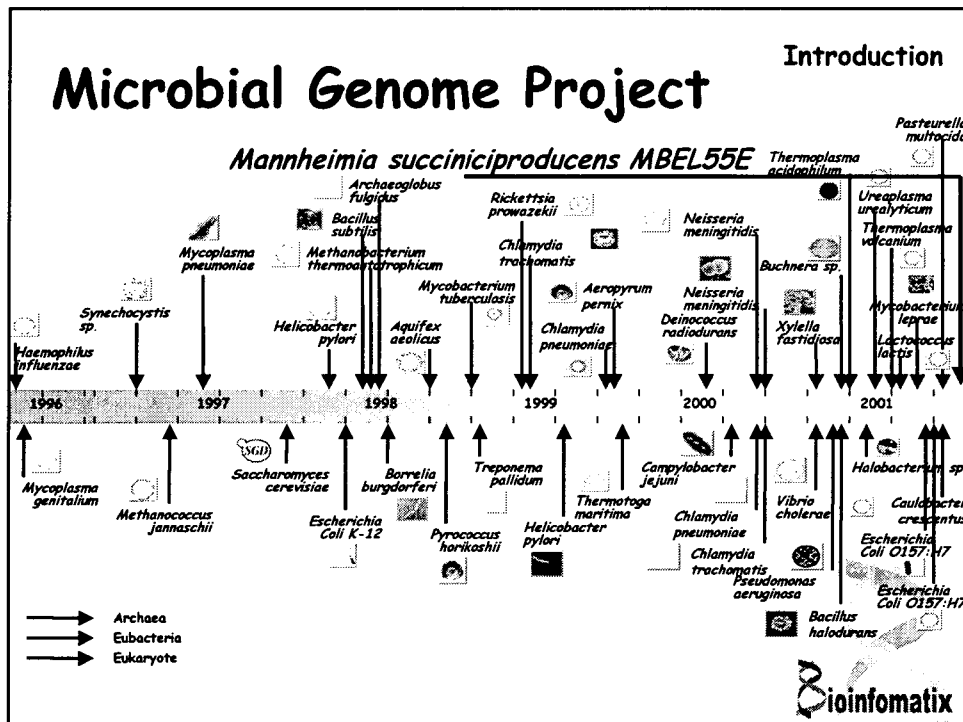
Contents

1. Introduction
2. *M. succiniciproducens* MBEL55E
3. Genome sequencing
4. Genome analysis
5. *In silico* metabolic network
6. Summary



Microbial Genome Project

Introduction



Completed genome project

Introduction

at June 14 2002

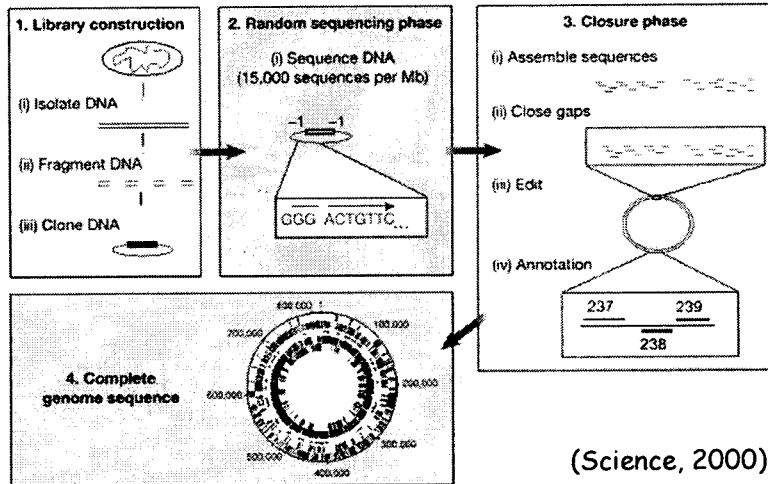
1. *Aeropyrum Pernix*
2. *Aquifex aeolicus*
3. *Archaeoglobus fulgidus*
4. *Bacillus halodurans*
5. *Bacillus subtilis*
6. *Borrelia burgdorferi*
7. *Buchnera sp.*
8. *Caulobacter crescentus*
9. *Campylobacter jejuni*
10. *Chlamydia pneumoniae CWL029*
11. *Chlamydia pneumoniae AR39*
12. *Chlamydia pneumoniae J138*
13. *Chlamydia trachomatis*
14. *Deinococcus radiodurans*
15. *Escherichia coli R1*
16. *Escherichia coli K-12*
17. *Escherichia coli O157*
18. *Haemophilus influenzae*
19. *Halobacterium sp.*
20. *Helicobacter pylori 26695*
21. *Helicobacter pylori J99*
22. *Methanobacterium thermoautotrophicum*
23. *Lactococcus lactis*
24. *Leishmania major Chr1*
25. *Methanococcus jannaschii*
26. *Mesorhizobium loti*
27. *Mycobacterium leprae*
28. *Mycobacterium tuberculosis*
29. *Mycoplasma genitalium*
30. *Mycoplasma pneumoniae*
31. *Mycoplasma pulmonis*
32. *Neisseria meningitidis*
33. *Neisseria meningitidis*
34. *Pasteurella multocida*
35. *Plasmodium falciparum Chr2*
36. *Plasmodium falciparum Chr3 and more*

■ Archae: 16 ■ Bacteria: 65 ■ Eucaryote: 12

Total 93



Whole-genome analysis



(Science, 2000)

Summary of features

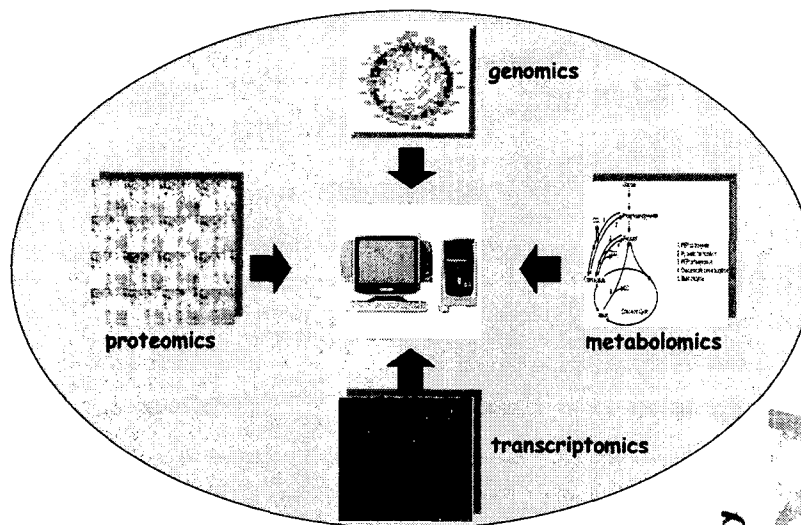
Organism	Genome size(Mbp)	Number of ORFs	Unknown	Unique
			Function	ORFs
<i>A. fulgidus</i>	2.18	2437	1315(54%)	641(26%)
<i>M. thermotauto.</i>	1.75	1855	1010(54%)	496(27%)
<i>M. jannaschi</i>	1.66	1749	1076(62%)	525(30%)
<i>P. horikoshi</i>	1.74	2061	859(42%)	453(22%)
<i>A. aeolicus</i>	1.50	1521	663(44%)	407(27%)
<i>B. subtilis</i>	4.20	4100	1722(42%)	1053(26%)
<i>B. burgdorferi</i>	1.44	1751	1132(65%)	682(39%)
<i>C. trachomatis</i>	1.04	894	290(32%)	255(29%)
<i>D. radiodurans</i>	3.28	3192	1715(54%)	1001(31%)
<i>E. coli</i>	4.60	4288	1632(38%)	1114(26%)
<i>H. influenzae</i>	1.83	1692	592(35%)	237(14%)
<i>H. pylori</i>	1.66	1657	744(45%)	539(33%)
<i>M. tuberculosis</i>	4.41	3924	1521(39%)	606(15%)
<i>M. genitalium</i>	0.58	470	173(37%)	7(2%)
<i>M. pneumoniae</i>	0.81	677	248(37%)	67(10%)
<i>Synechocystis</i> sp.	3.57	3168	2384(75%)	1426(45%)
<i>T. martima</i>	1.86	1877	863(46%)	373(26%)
<i>T. pallidum</i>	1.14	1040	461(44%)	280(27%)
<i>R. prowazekii</i>	1.10	834	48(12%)	207(25%)
<i>C. pneumoniae</i>	1.23	1073	437(40%)	186(17%)
<i>A. pernix</i>	1.67	2694	2061(76%)	1538(57%)
<i>L. lactis</i>	2.35	1495	398(27%)	83(6%)
	45.6	44449	20726(47%)	11924(27%)

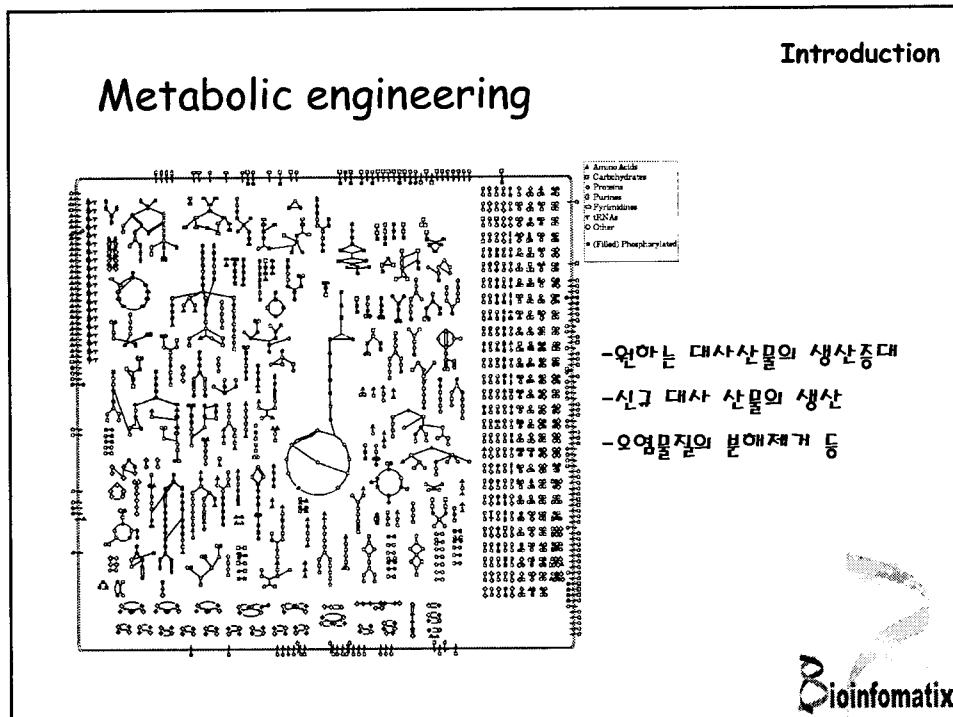
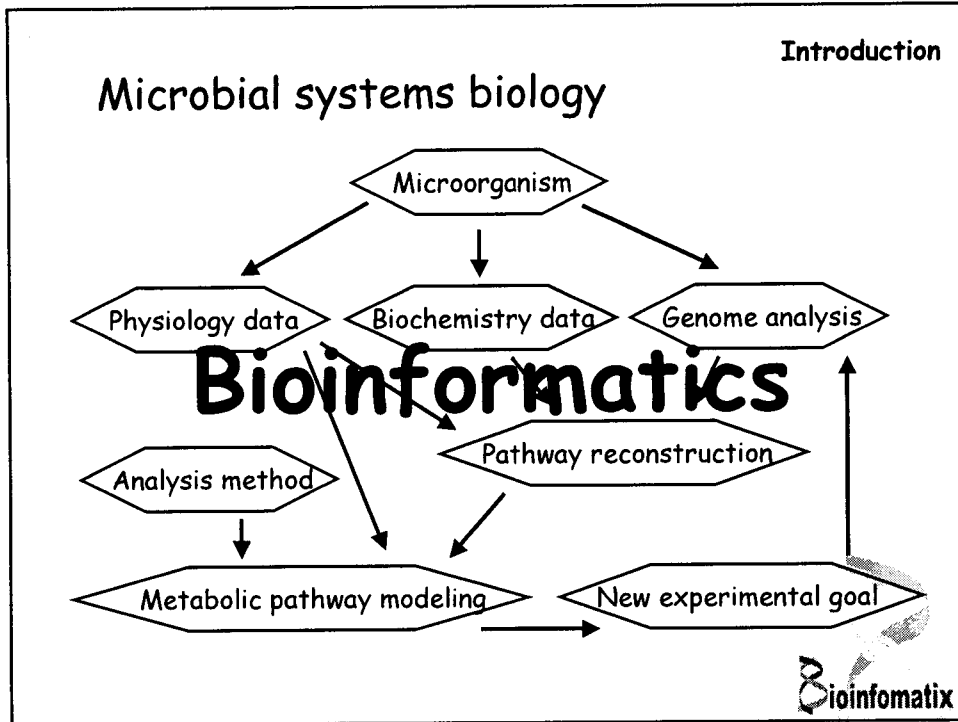
Post-genome Era

Genomics, proteomics, metabolomics and bioinformatics are accelerating the successful development of metabolic engineering strategies at system level

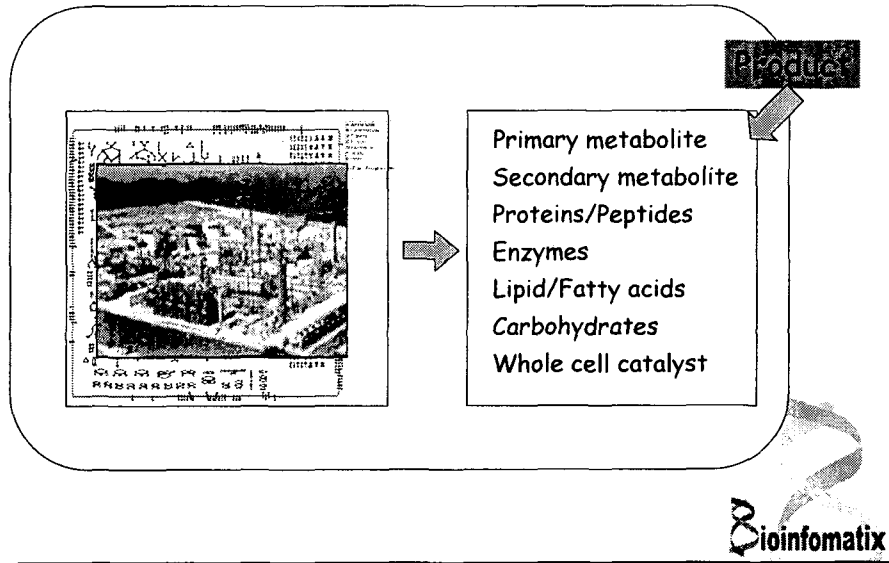


Metabolic pathway based on systems biology





Microorganisms as chemical factories



Industrial microorganisms

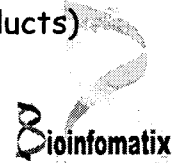
- Streptomyces* ----- antibiotics and other secondary metabolites
- Bacillus* ----- enzyme
- Corynebacterium* ----- amino acid
- Thermophilic, other archaeobacteria* ---- novel enzyme
- Pseudomonas* ----- bioconversion, bioremediation
- Clostridium* ----- solvent production, cellulase system
- Deinococcus* ----- radiation resistance
- Rhodobacter* ----- photosynthesis



M. succiniciproducens

Fermentative production

- ❖ Fermentative production → Green technology
- ❖ The industrial scale production
 - ⇒ requires reduction in its production cost.
- ❖ One possible way to reduce the cost of the fermentation process
 - ⇒ use renewable carbon sources
(agricultural and dairy waste products)



M. succiniciproducens

Succinic acid synthesis

Chemical process

n-Butane ⇒ maleic anhydride ⇒ maleic acid ⇒ Succinic acid

Fermentation

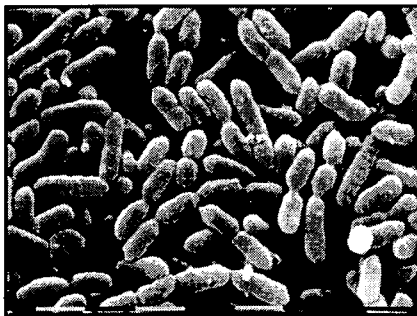
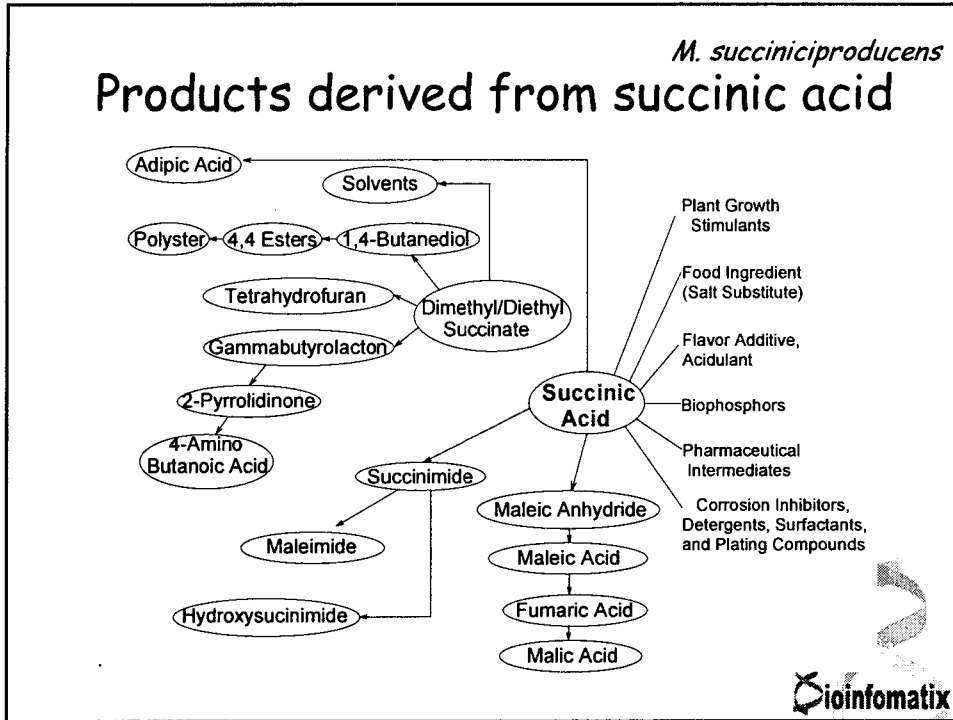
Renewable biomass ⇒ (microorganisms) ⇒ Succinic acid

Bioconversion

Too expensive

Fumaric acid ⇒ (microorganisms or enzymes) ⇒ Succinic acid

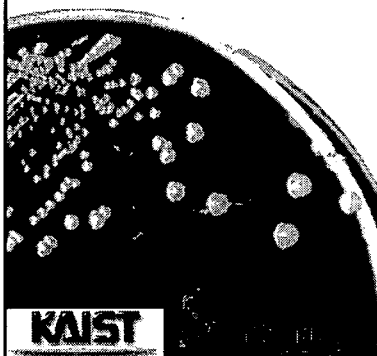




Isolation of new succinic acid producer

Mannheimia succiniciproducens MBEL55E

- Isolated from bovine rumen
 - Non-motile
 - Non-spore-forming
 - Mesophilic
 - Capnophilic
 - Gram-negative



Bioinformatix

M. succiniciproducens

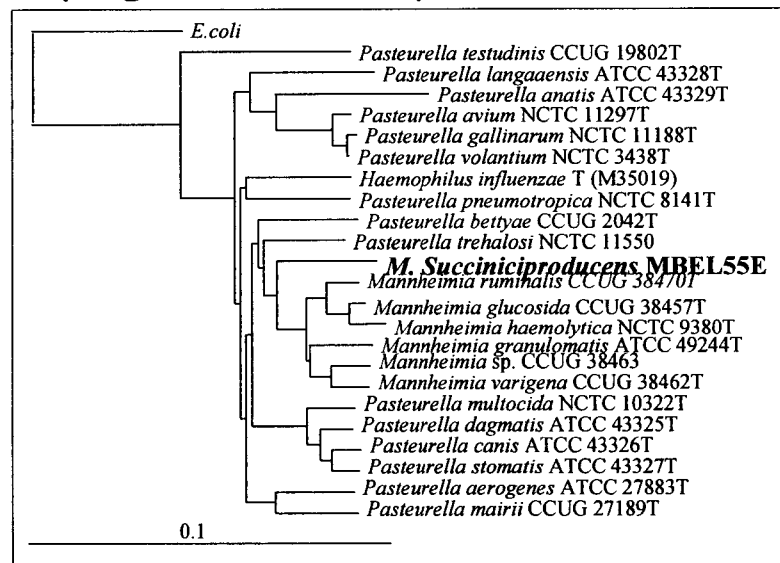
16s rDNA similarity

Species	Accession Number	Similarity (%)
<i>Mannheimia succiniciproducens</i> MBEL55E (KCTC 0769BP)	AY029193	100
<i>Mannheimia varigena</i> CCUG 38462 ^T	AF053893	95.1
<i>Mannheimia granulomatis</i> ATCC 49244 ^T	AF053902	94.9
<i>Mannheimia ruminalis</i> CCUG 38470 ^T	AF053900	94.7
<i>Mannheimia glucosida</i> CCUG 38457 ^T	AF053889	94.7
<i>Pasteurella trehalosi</i> NCTC 11550	U57073	94.7
<i>Pasteurella dagmatis</i> ATCC 43325 ^T	M75051	94.3
<i>Mannheimia haemolytica</i> NCTC 9380 ^T	AF060699	94.1
<i>Pasteurella avium</i> NCTC 11297 ^T	M75058	94.0
<i>Pasteurella aerogenes</i> ATCC 27883 ^T	M75048	94.0
<i>Pasteurella bettyae</i> CCUG 2042 ^T	L06088	94.0
<i>Pasteurella volantium</i> NCTC 3438 ^T	M75070	93.8
<i>Pasteurella mairii</i> CCUG 27189 ^T	AF024532	93.8
<i>Pasteurella multocida</i> NCTC 10322 ^T	M35018	93.8
<i>Haemophilus influenzae</i> ^T	M35019	92.5
<i>Vibrio cholerae</i> ATCC 14035 ^T	Z21856	87.8
<i>Escherichia coli</i>	J01695	86.6

Bioinformatic

M. succiniciproducens

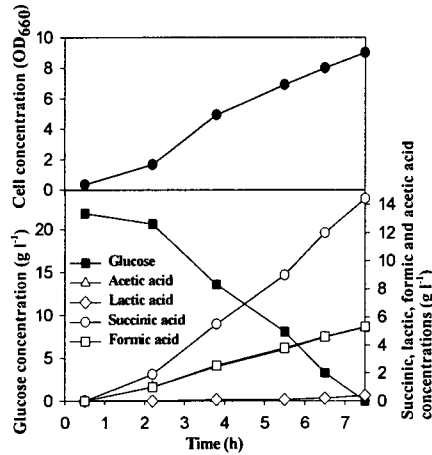
Phylogenetic tree position



Bioinformatic

M. succiniciproducens

Anaerobic batch culture of *Mannheimia succiniciproducens* MBEL55E



- Substrate
 - 20 g/L glucose
- Product
 - 14 g/L succinic acid
 - 5.5 g/L of acetic and formic acids

KAIST Bioinformatix

Completed for the first time in the world



바이오인포메틱스(주)





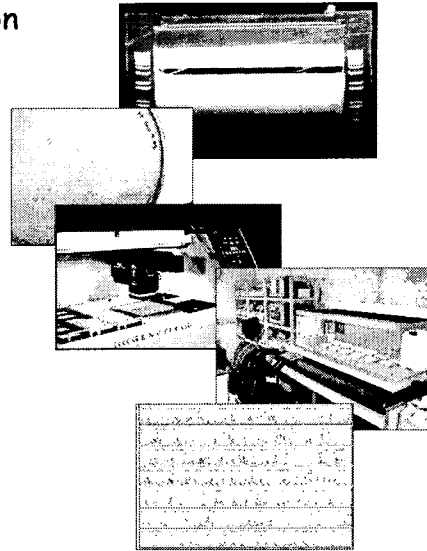
Best Partner in Genomic Research



Bioinformatix

Overview of Sequencing strategy

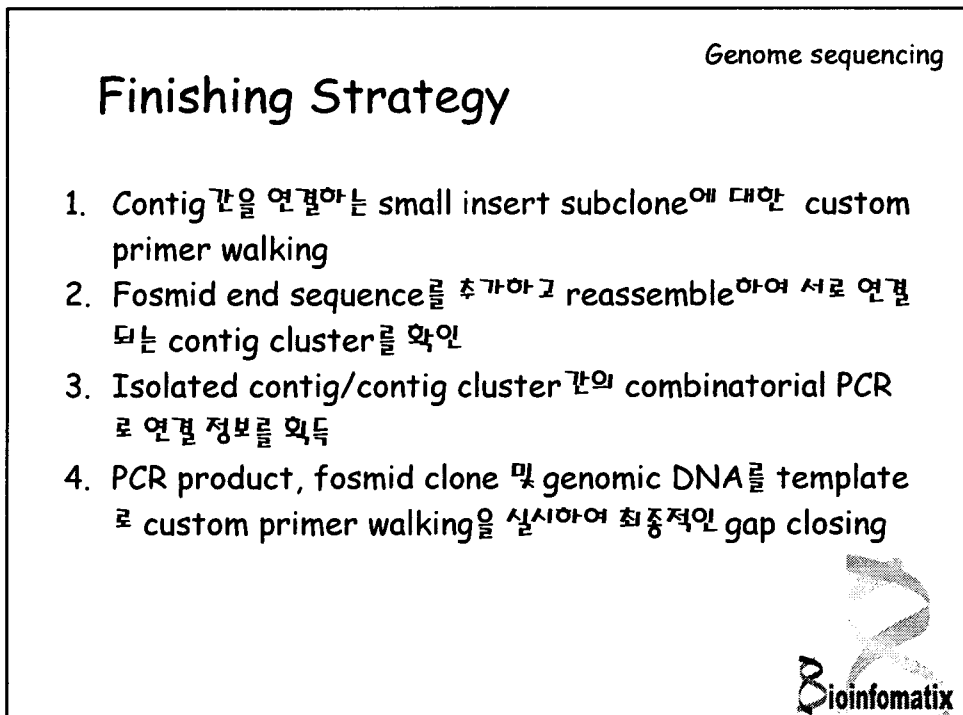
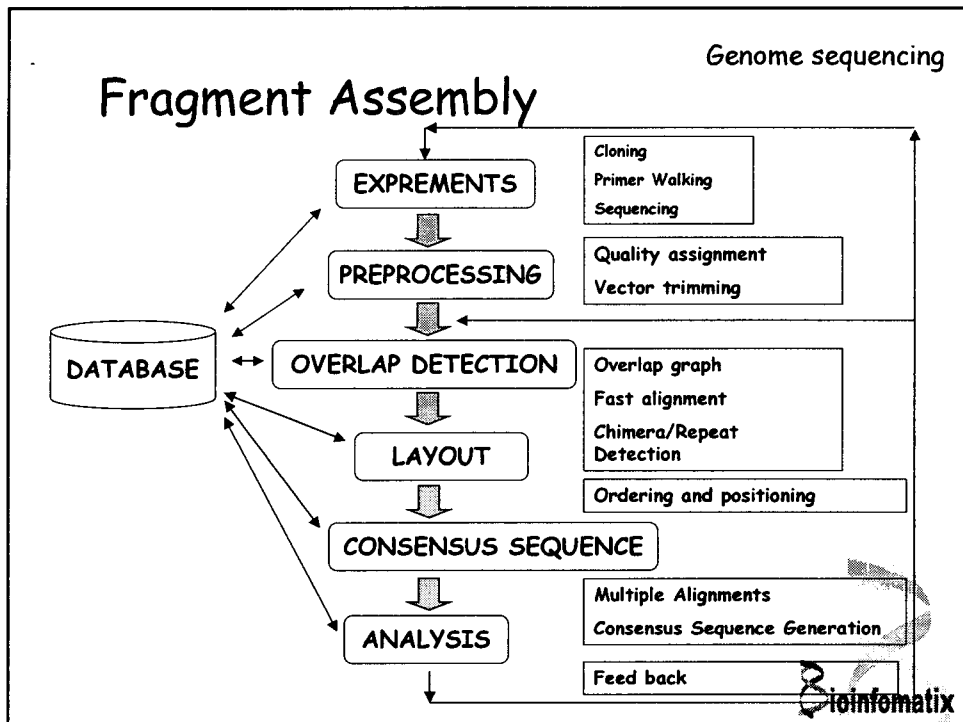
1. Shotgun library construction
 1. Genomic DNA preparation
 2. Ligation
 3. Transformation
2. High throughput Sequencing
 1. Recombinant clone culture
 2. DNA preparation
 3. Clone end sequencing
 4. Quality assessment
3. Contig assembly
4. Finishing
 1. Contig ordering & gap closing
 2. Manual editing
 3. Low quality region confirm



Input Data Quality (total 38,911 reads)

1. 36,820 reads from small insert library (~3.5kb)
 - 22,599,739 bp(trimmed subsequence)
 - : avg 613.8 bp/read
 - about 20 bp of vector sequence per read
2. 530 reads from fosmid library (~38.5kb)
3. 1,561 custom primer walk reads





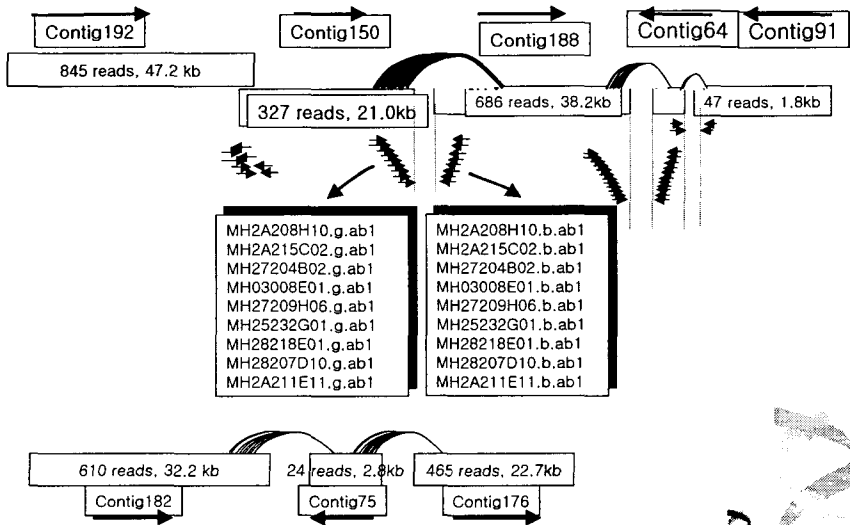
Contig Assembly Results

1. After shotgun phase(2001. 5.)
 - 총 195개의 Contig (34,820 subclone reads)
 - 98.47% coverage (P.mul 기준)
 - 83개의 scaffold로 구성

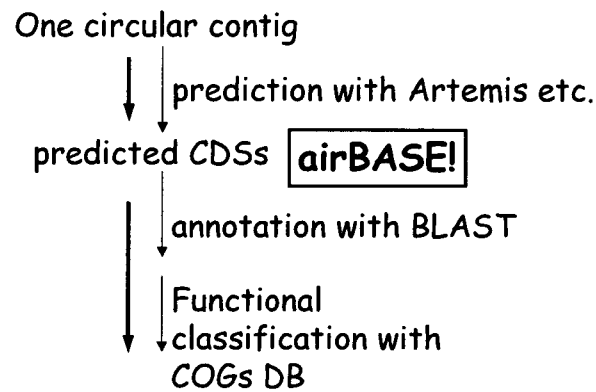
2. After finishing (2001.8.)
 - One circular contig (38,911 reads)
 - 2,314,078 bp



Example of contig assembly



CDSs Prediction and Annotation



About airBASE!

- What's airBASE!™?
 - Integrated genome annotation system
 - Annotated information retrieval system
 - Automatic database construction
 - Client/server environment
 - Simple but powerful
 - easy to use, familiar graphic user interface
 - fast and various search

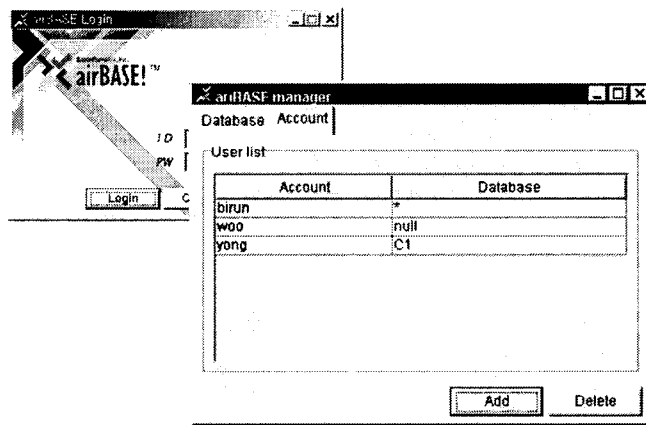


Functional supports

- **System Management**
 - Login system
 - User management
- **Automatic annotation and database construction**
- **Data search and visualization**
 - Gene view
 - Overview
 - Keyword search
 - Pathway search
 - Function search



System management



Automatic annotation

The screenshot displays the 'Automatic Annotation' software interface. On the left, a vertical sidebar lists five steps: Step 1 (Input Whole Genome), Step 2 (Running Artemis), Step 3 (Input Artemis result), Step 4 (Input Database Name), and Step 5 (Running Automatic Annotation). The main window shows a detailed log of the process, including parameters like 'Total space allocated: 13,508 Mbytes' and 'Total # pairs: 4000'. Below the log, a sequence alignment is visible with nucleotide bases (A, C, G, T) and a corresponding amino acid sequence (I, L, V, A, T, D, S, A, G, L, L, T, M, I, S, A, K, E, A, A, I, A, E, L).



Visualization and search

The screenshot shows the 'Database Search' software interface. It features a 'Select database' dropdown menu, a search input field, and a 'Search' button. A search results table is displayed with columns for 'No.', 'Description', 'Length', 'Score', and 'Type'. Below the table, a list of biological functions is shown, including 'BASIC TRANSCRIPTION FACTORS', 'TRANSCRIPTIONAL REGULATORS', 'BASAL REPLICATION MACHINES', and 'DNA POLYMERASE II SUBUNITS'. The Bioinformatics logo is visible in the bottom right corner.

in silico metabolic network

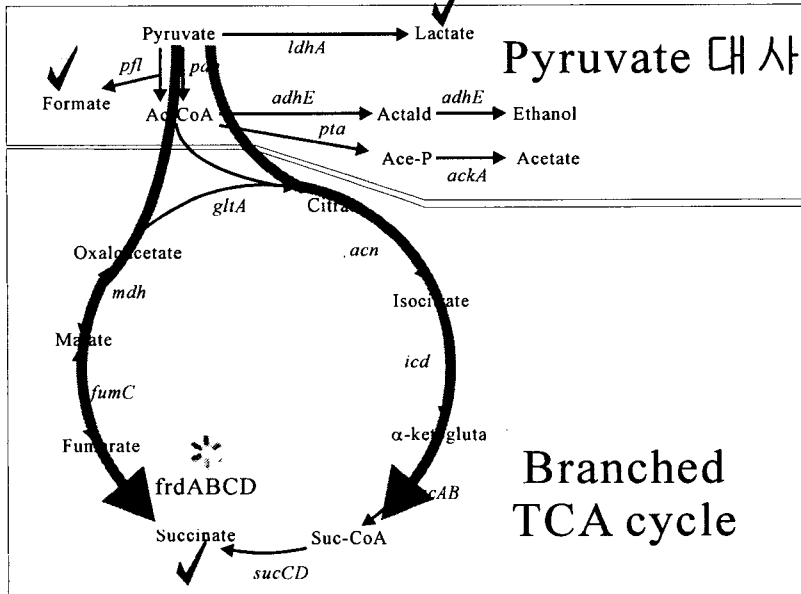
In silico metabolic network of *M.s.*

A17	B	C	D	E
32	The Tricarboxylic Acid Cycle			
33	28 <i>glfA</i> Genome	Citrate synthase		ACCOA + OA -> COA + CIT
34	29 <i>acnA</i> Genome	Aconitase A		CIT -> ICIT
35	30 <i>icdA</i> Genome	Isocitrate dehydrogenase		ICIT + NADP -> CO2 + NADPH + AK
36	31 <i>sucAB, lpcdA</i> <i>In silico</i>	2-Ketoglutarate dehydrogenase		AKG + NAD + COA -> CO2 + NADH + AK
37	32 <i>sucCD</i> Genome	Succinyl-CoA synthetase		SUCCOA + ADP + PI -> ATP + COA +
38	33 <i>sucABC</i> Genome	Succinate dehydrogenase		SUCC + FAD -> FADH + FUM
39	34 <i>frdABCD</i> Genome	Fumarate reductase		FUM + FADH -> SUCC + FAD
40	35 <i>fumA</i> Genome	Fumarase A		FUM -> MAL
41	36 <i>mdh</i> Genome	Malate dehydrogenase		MAL + NAD -> NADH + OA
42				
43	Pyruvate Metabolism			
44	37 <i>ldhA</i> Genome	D-Lactate dehydrogenase 2		PYR + NADH -> NAD + LAC
45	38 <i>adhE</i> <i>In silico</i>	Acetaldehyde dehydrogenase		ACCOA + 2 NADH -> ETH + 2 NAD +
46	39 <i>pta</i> Genome	Pyruvate formate lyase 1		PYR + COA -> ACCOA + FOR
47	40 <i>pta</i> Genome	Phosphotransacetylase		ACCOA + PI -> ACTP + COA
48	41 <i>ackA</i> Genome	Acetate kinase A		ACTP + ADP -> ATP + AC
49	42 <i>acs</i> Genome	Acetyl-CoA synthetase		ATP + AC + COA -> AMP + PPI + ACC
50				
51	Anaplerotic Reactions			
52	43 <i>pckA</i> Genome	Phosphoenolpyruvate carboxylase		OA + ATP -> PEP + CO2 + ADP
53	44 <i>ppc</i> Genome	Phosphoenolpyruvate carboxylase		PEP + CO2 -> OA + PI



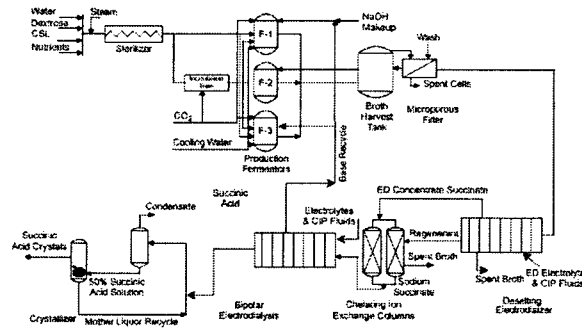
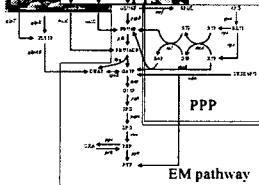
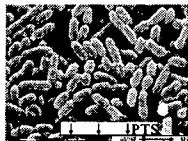
TCA cycle & pyruvate metabolism

KAIST Bioinformatic
바이오인포매틱스(주)



After the genome project and metabolic engineering...

→ Development of Super-Succinic acid producer



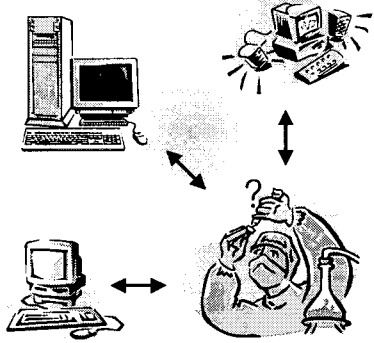
Acknowledgement



한국과학기술원 이상엽 교수팀, (주)제노텍, 한국생명공학 연구원 어철규 실장팀, 그리고 바이오인포메틱스(주) 참여 연구자 분들께 감사드립니다.



Looking for Partners !!!



135-818 서울시 강남구 논현동 81-7

바이오인포매틱스(주) 부설 연구소

Tel (02)3445-4533

Fax (02)3445-1281

Email : lab@bioinformatix.com

Thanks for your attention

