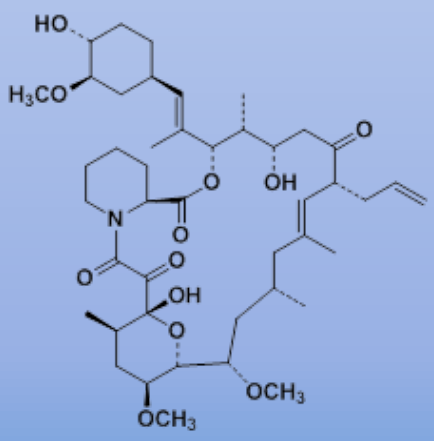
N²PC Next generation Natural Products Chemistry

Technology Research Association for Next-generation Natural Products Chemistry

Technical Information



What Is the Technology Research Association for Next-generation Natural Products Chemistry?

Having a wide range of bioactivity and great structural diversity, natural compounds produced by microorganisms are highly promising as candidate compounds for new drugs. In addition, natural-compound chemistry is a discipline in which Japan has a global presence in both basic research and industrial applications, with its long-developed unique technical resources and know-how. However, natural-compound libraries could not emerge and evolve without many processes, including the management and culture of microbial strains and the isolation and synthesis of hit compounds, which often involve painstaking treatments. Their handling as such necessitates special experience and technical mastery. Unfortunately, however, there has recently been a tendency for Japanese companies to reconsider (reduce) their drug discovery research involving the use of natural compounds.

Against this background, some companies having extensive libraries jointly founded our association in 2011 to promote cooperative research in drug discovery for pharmaceutical and other companies. Since libraries developed independently by companies are precious properties, they have traditionally been limited to in-house usage. In contrast, this project promotes the mutual use of such libraries among the member companies of the Association and Academia, with unified maintenance and control by the Association. This mutual use system for natural-compound libraries is expected to maximize the mutual use of libraries owned by individual companies to the benefit of both the providers and users of the libraries, which in turn can lead to the discovery of more drug candidate compounds, thus contributing to the development of natural-substance-based drug discovery in Japan.

Overview of the Technical Resources of the Technology Research Association for Next-generation Natural Products Chemistry

Kazuo Shin-ya, Project Leader

Natural compounds produced by microorganisms seem to represent an ideal resource for libraries of intermediate molecular compounds. Their total synthesis, however, is often extremely difficult, although not completely impossible by organic synthesis. Furthermore, their derivatization has been practically unfeasible because of necessity for total synthesis from starting materials. In this situation, there has been a long history of attempts to obtain their derivatives by modifying parent cores through manipulation of biosynthesis genes. However, for representative natural compounds in wide use as clinical drugs, such as macrolides and cyclopeptides, it has not been possible to achieve purposeful genetic modifications using conventional gene manipulation technology based on the use of homologous recombination and restriction endonucleases, because their biosynthesis genes comprise giant gene clusters exceeding 100 kb in length, and involve extremely high homology and sequence repeats. In addition, these natural compounds have generally only been produced through bacteria-specific production; therefore, it has been a still higher hurdle to achieve accurate genetic manipulation in the presence of a great many biosynthesis genes other than the desired ones.

Since 2011 we have been engaged in technical development through a number of projects sponsored by NEDO, METI, and AMED. Thereby, we have succeeded in establishing technologies for accurate genetic analysis and preparation of libraries of BACs exceeding 200 kb in length. To date, we have identified many biosynthesis genes of intermediate-molecular weight natural compounds exceeding 100 kb, acquired a great deal of information on their sequences, and established heterologous expression technologies with the use of such genes. Combining this technology and domain-swapping technology, we are working to develop innovative technology for genetically modifying the biosynthesis genes of intermediate-molecular weight natural compounds, thus enabling purposeful modifications of the backbones of such compounds.

On the other hand, as seen in the clinical development of Mevalotin® and ivermectin, drug effects and metabolism can be improved dramatically simply by reducing the only hydroxyl group or double-bond. In organic synthesis, scaffold-free synthesis and position-specific reactions are quite difficult. Hence, our technical development focuses on a library of enzymes that catalyze reactions that could not be carried out by organic synthesis, and that enable us to modify high-molecular weight compounds such as macrolides. We are now working to modify various compounds using the library, and to obtain their derivatives by an organic synthesis approach using scaffolds. Also engaged in the development of expression vectors for the activities of various enzymes, we have been successful in the expression of the activities of P450 and other enzymes in actinomycetes. While conventional modification enzymes are for the most part modifiers of low-molecular weight compounds such as industrial raw materials, we have found enzymes capable of modifying some intermediate-molecular weight natural compounds. Using docking data for such enzymes and compounds and other information, we are also working to develop simulation technologies, including enzyme specificity modifications.

We hope you will make the best use of these technologies via the Technology Research Association for Next-generation Natural Product Chemistry to facilitate your business activities, including drug discovery and agrochemical development from natural compounds.

Accurate Gene Analysis

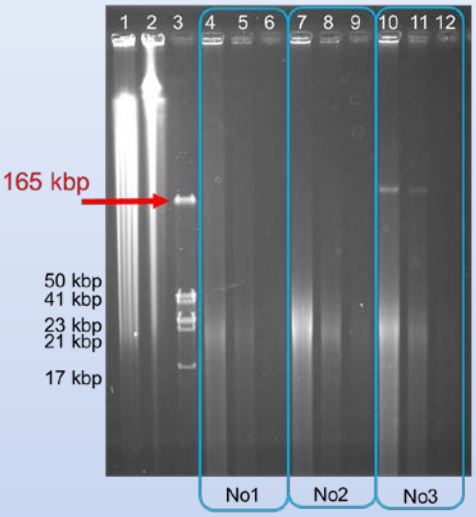
We have succeeded in establishing a technical scheme for accurate analysis of complexes such as actinomycete genome by combining short-lead sequencing with Miseq and long-lead sequencing with PacBio.

Miseq

PacBio

low-molecular weight genome

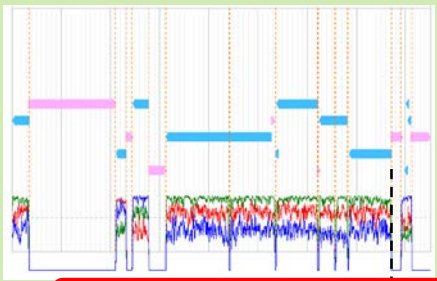
	name		DNA濃度 (ng/μL)	総ゲノム泳動量(ng)
1	con-1	EM52	300	3000
2	con-2	EM52	438	4380
3	マーカー	7-GT		
4				35300
5	No. 1		7060	7060
6				1412
7				24350
8	No. 2		4870	4870
9				974
10				35775
11	No.3		7155	7155
12				1431



Preparation of high-molecular weight genome is essential for successful sequencing

Example sequencing results for type I polyketide biosynthesis gene consisting of a high degree of repeat sequences

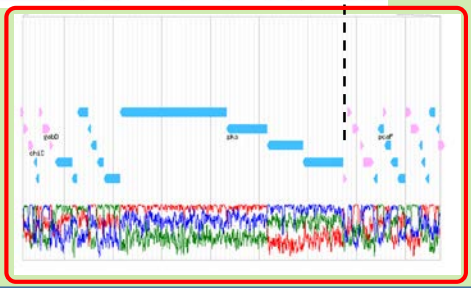
Analysis results from another research organization



Number of scaffolds: 262

Region plotting in accordance with decoded sequence data
Red line: fragmentation site.
Blue band: analyzed region
Pink band: unanalyzed region

Results of analyses by the Association



Number of scaffolds: 2

Blue band: right-to-left sequence
Pink band: left-to-right sequence

- ✓ A combination of high-quality genome and our own sequence data processing enables analyses of nearly-full-length genome of actinomycetes.
- ✓ Permits analyses of GC-rich genome regions with high-degree repeat sequences, such as actinomycete biosynthesis genes.

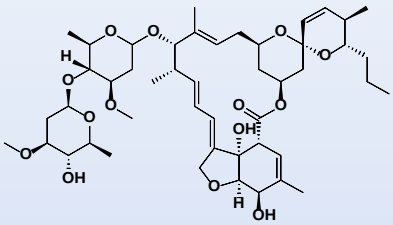
In humans and other mammals, promoter domains have high GC contents, and many gene regions responsible for multiple diseases are known to have high-degree repeat sequences. The actinomycete genome analysis technology developed through this project and the technical abilities compiled on the basis of the results of many analyses in combination are expected to produce valuable data of paramount importance for personalized medicine in the near future.

Acceptable contract operations
Sequence analysis

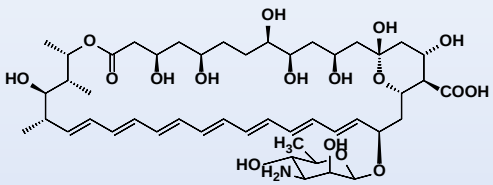
Preparation Technology for a Library of BACs* Longer Than 200 kb

*BAC: Bacterial artificial chromosome, artificial *Escherichia coli* chromosomes

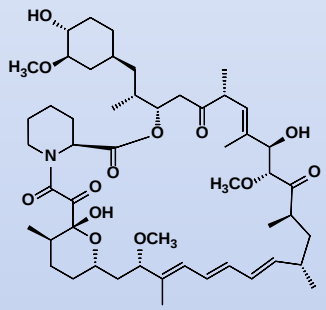
*BAC: Bacterial artificial chromosome. Gene fragments up to 300 kbp can be cloned (the largest 266-kbp clone obtained).



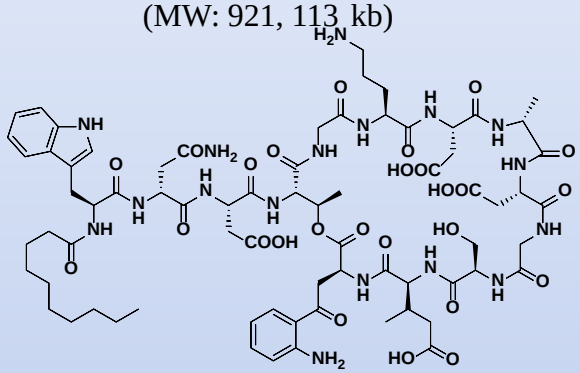
Avermectin
(MW: 858, 81 kb)



Amphotericin B
(MW: 921, 113 kb)

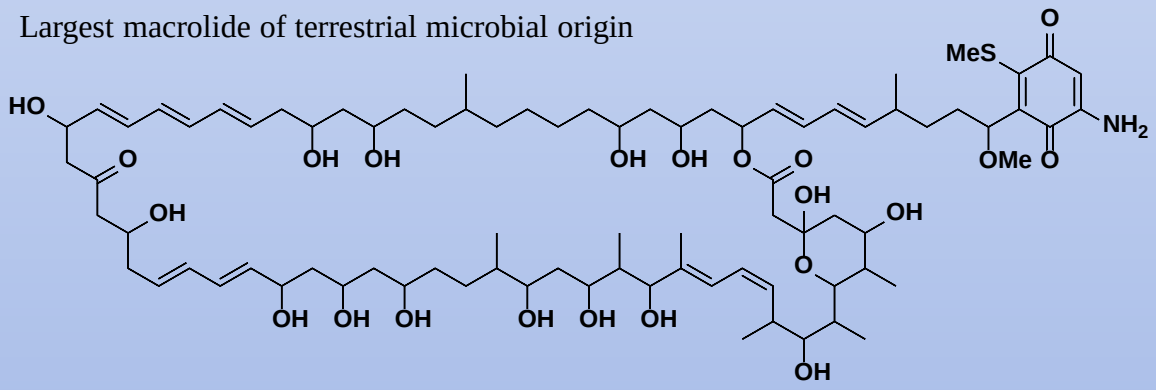


Rapamycin
(MW: 913 107 kb)



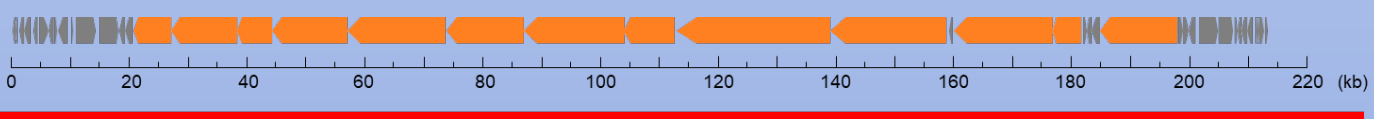
Daptomycin
(MW: 1620, 128 kb)

Largest macrolide of terrestrial microbial origin



Cluster size: 213.9 kbp
Insert length: 232.2 kbp

Quinolidomicin
$C_{83}H_{132}O_{23}S$, MW: 1529



232.2 kb

Reference: T. Hashimoto, et al. Biosynthesis of quinolidomicin, the largest known macrolide of terrestrial origin: Identification and heterologous expression of a biosynthetic gene cluster over 200 kb, *Org. Lett.*, **20** (24), 7996-7999 (2018).

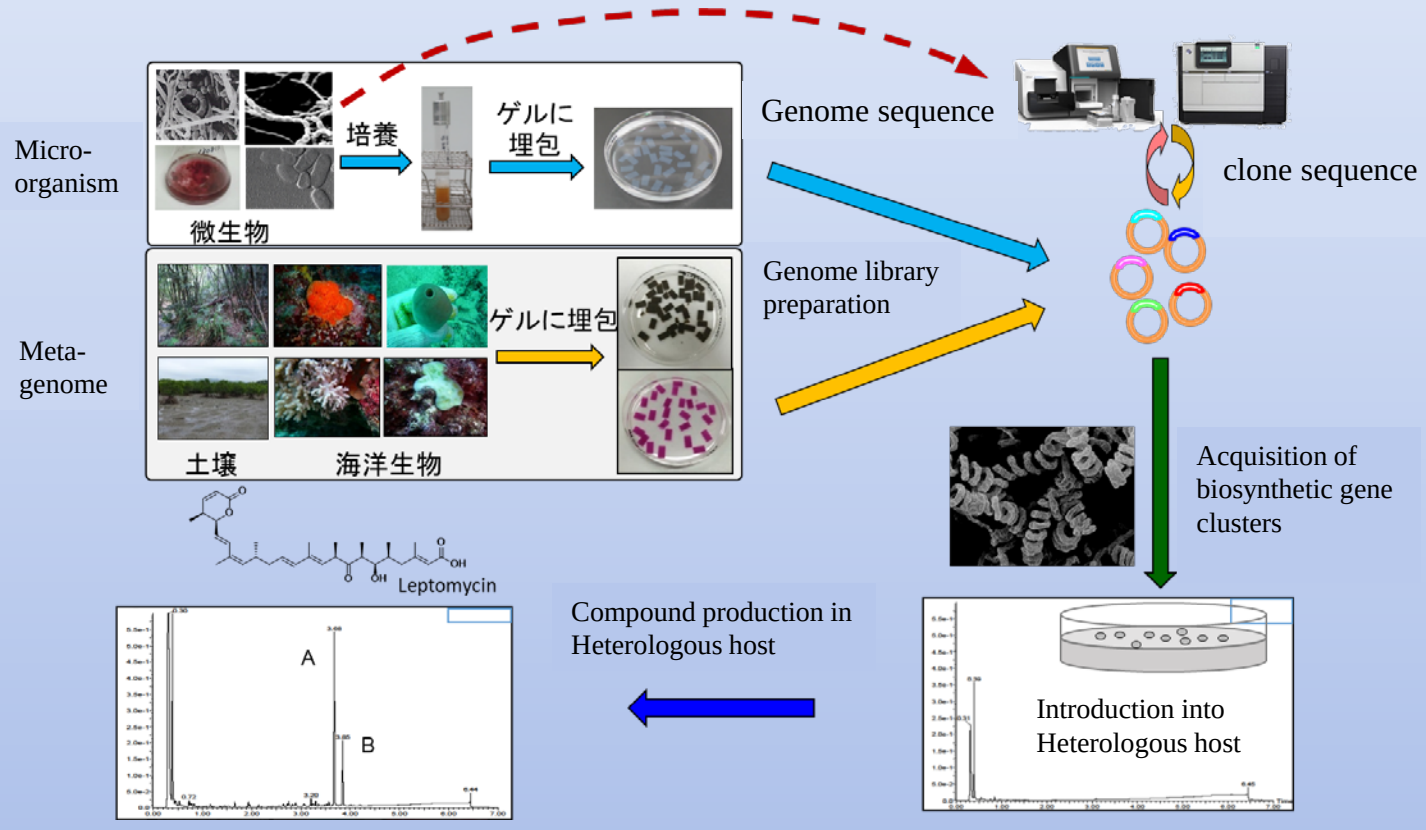
Acceptable contract operations

Contract preparation of BAC libraries

Developing a BAC library from bacterial producers of useful natural compounds on a contract basis
Because cloning is achieved using BAC, long genes can be cloned as they are.

Production of Secondary Microbial Metabolites by Heterologous Expression

- ✓ Culturable microorganisms may be cultured and analyzed for nearly-full-length genomes. Even for meta-genomes, long genome libraries can be prepared using a method for high-quality genome preparation.
- ✓ After acquiring the desired biosynthesis gene cluster via sequencing, the cluster may be introduced into a highly productive host for heterologous expression to attain higher productivity than with the original producer.



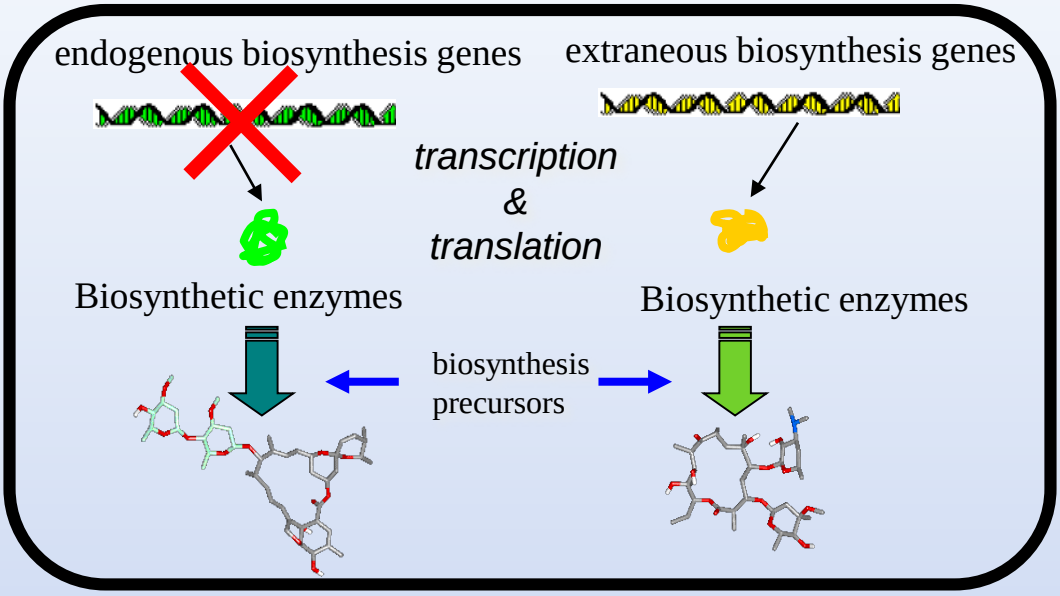
Reference: Shin-ya K, Ikeda H. Production of natural compounds by heterologous expression (in Japanese). *Kagaku-to-Seibutsu*, Vol. 52 (9), 616-621 (2014).

Acceptable contract operations

Production of secondary microbial metabolites by heterologous expression

1. Microbial genome decoding
2. Preparation of genome libraries using BAC vectors
3. Cloning of desired biosynthesis gene
4. Production by heterologous expression in clean host strains
5. Core modification by biosynthesis gene modification

Characterization of Host Strains for Heterologous Expression

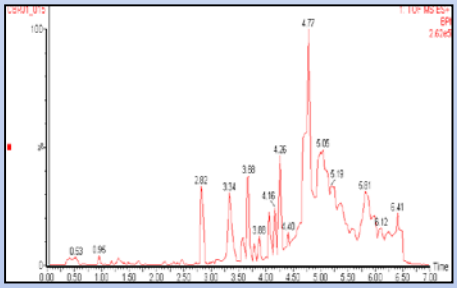


"Clean hosts" for heterologous expression of secondary metabolites

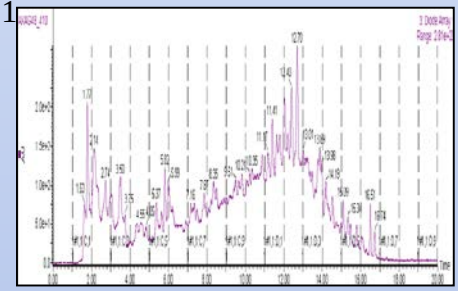
Since endogenous biosynthesis genes that compete for biosynthesis precursors have been removed, it is easy to express extraneous biosynthesis genes.

Culture broths of microorganisms collected in nature (wild strains) contain numerous compounds.

Example 1

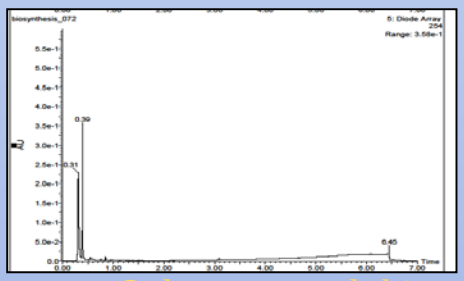


Example 1

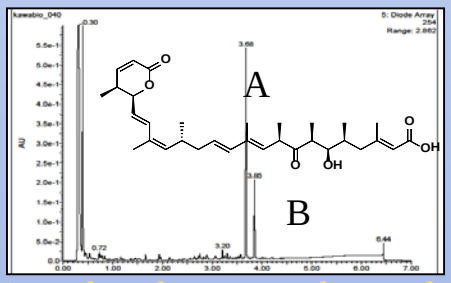


Example host strains used by the Association

Host strain (not transfected)



Host strain (transfected)



Only compounds biosynthesized by the introduced gene are detected.

Reference: H. Ikeda, K. Shin-ya, and S. Omura. Genome mining of the Streptomyces avermitilis genome and development of genome-minimized hosts for heterologous expression of biosynthetic gene clusters, J. Ind. Microbiol. Biotechnol., 41, 233 (2014).

Acceptable contract operations

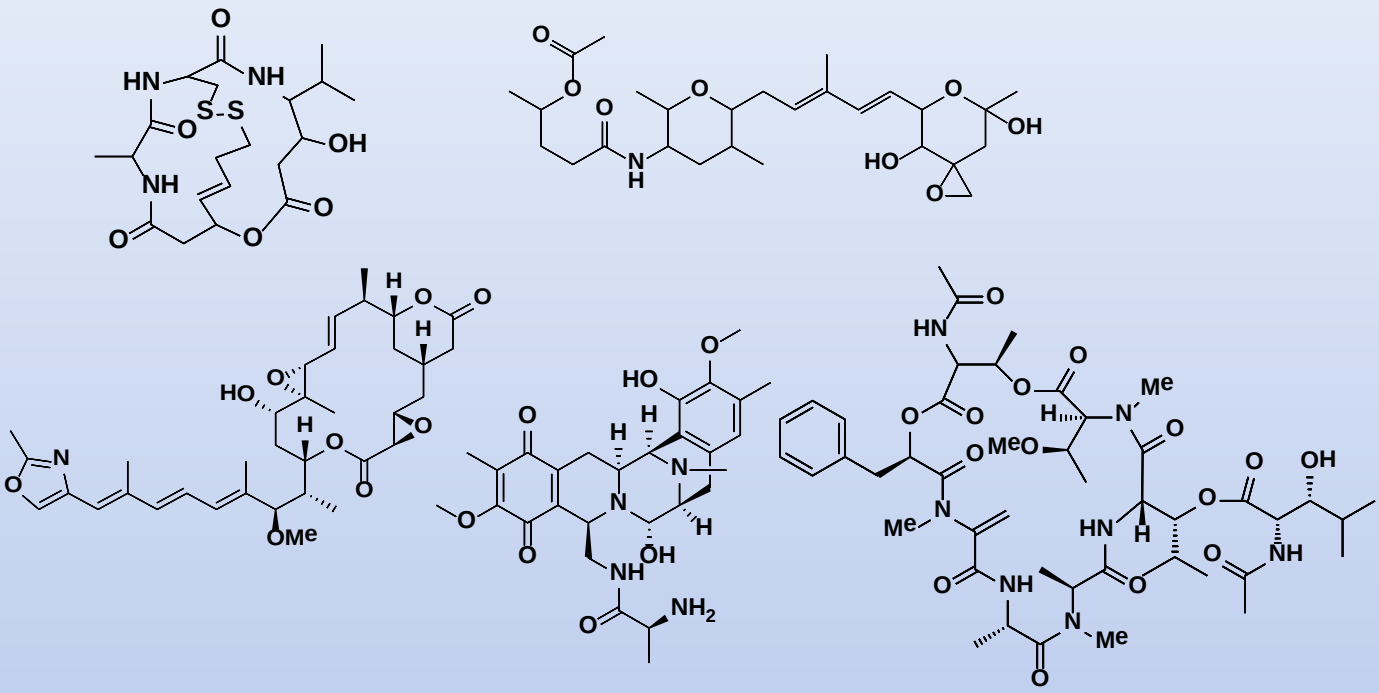
Production of secondary microbial metabolites by heterologous expression

1. Microbial genome decoding
2. Preparation of genome libraries using BAC vectors
3. Cloning of desired biosynthesis gene
4. Production by heterologous expression in clean host strains
5. Core modification by biosynthesis gene modification

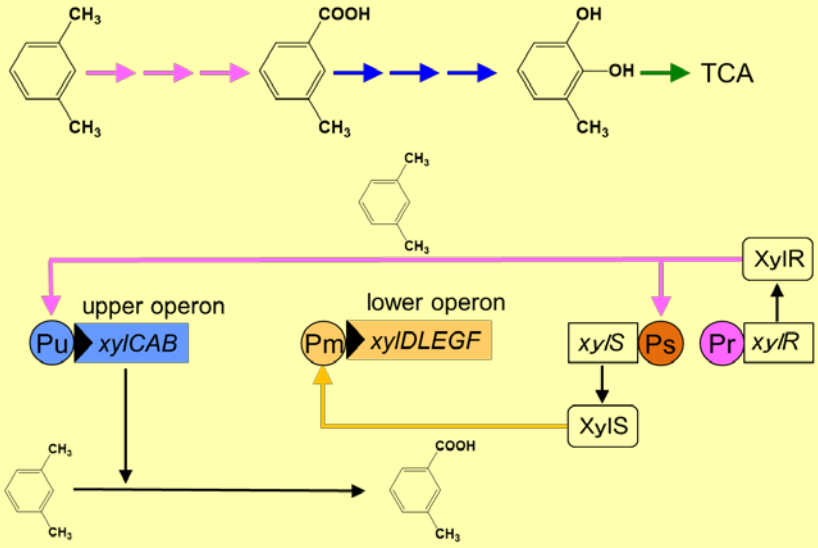
Heterologous Expression System for *Pseudomonas* Bacteria

✓ Since hosts with lower GC contents than GC-rich actinomycete hosts are suitable for biosynthesis genes of non-actinomecetous organisms with low GC contents, we have also developed a *Pseudomonas* host for heterologous expression.

Heterologous expression using *Pseudomonas putida* / attB1 strain



Various *Pseudomonas* promoters that include a conditional expression system

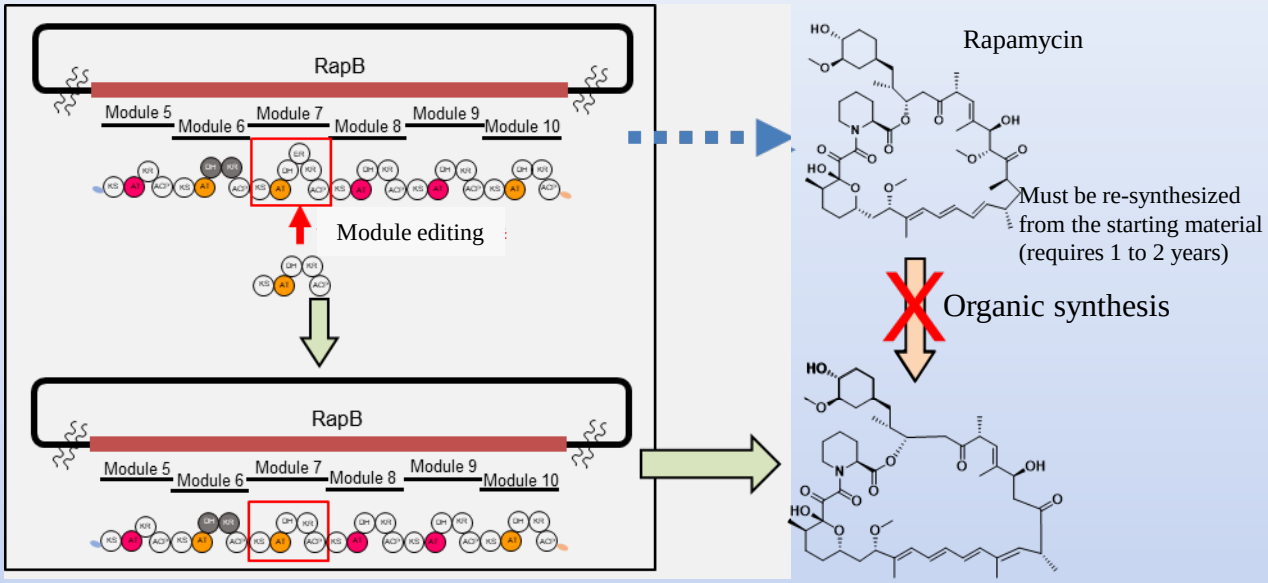


<u>Acceptable contract operations</u>

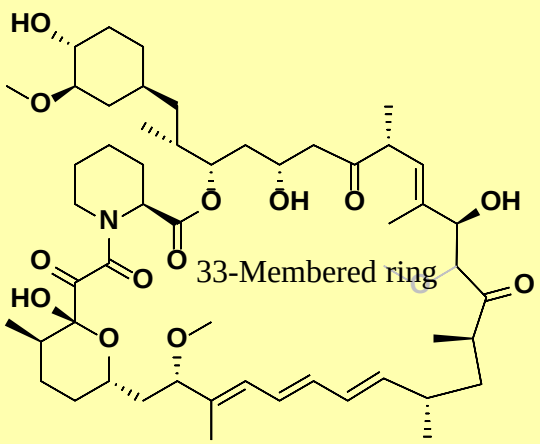
Production of secondary microbial metabolites by heterologous expression

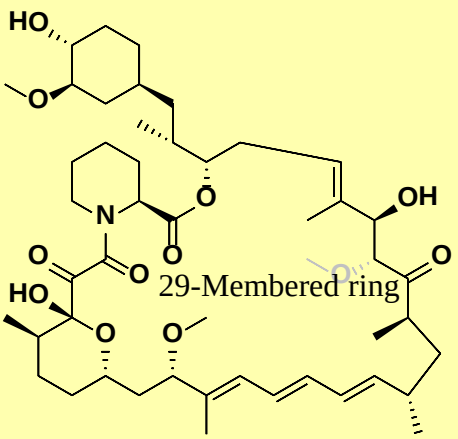
1. Microbial genome decoding
2. Preparation of genome libraries using BAC vectors
3. Cloning of desired biosynthesis gene
4. Production by heterologous expression in clean host strains
5. Core modification by biosynthesis gene modification

Backbone Modification Technology by Biosynthesis Gene Modification



Example compounds with modified rapamycin cores





Reference: Production method for core-modified compounds, patent pending.

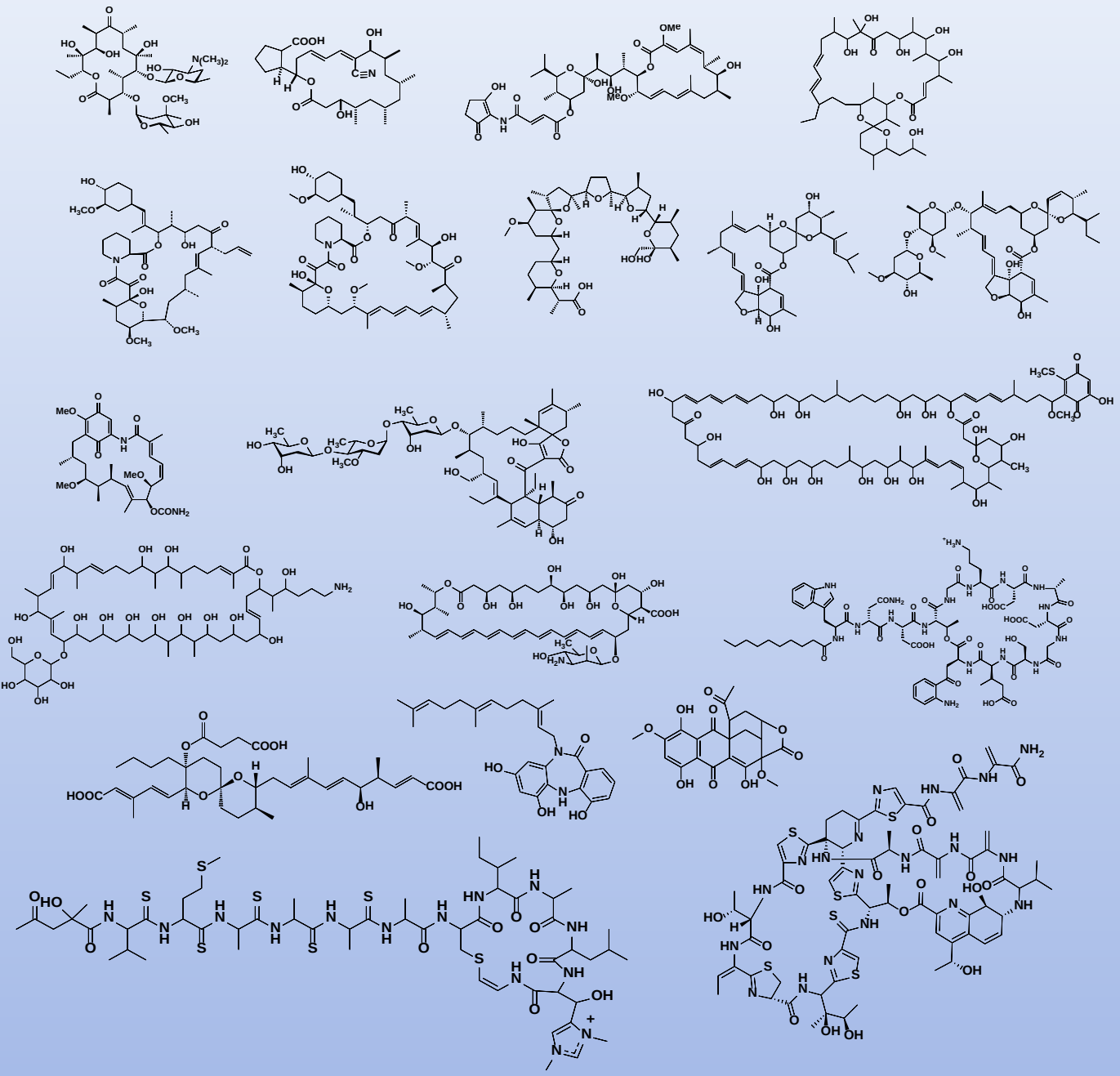
Acceptable contract operations

Module editing

Preparation of analogues of useful natural compounds

Biosynthesis Gene Clusters, Expression Vectors, etc.

BAC clones containing a giant biosynthesis gene cluster (examples)



Vectors optimized for actinomycete expression (P450 and other enzymes, transcription control factors, etc.)

<u>Acceptable contract operations</u>

Production of secondary microbial metabolites by heterologous expression

1. Microbial genome decoding
2. Preparation of genome libraries using BAC vectors
3. Cloning of desired biosynthesis gene
4. Production by heterologous expression in clean host strains
5. Core modification by biosynthesis gene modification

History

April 2006	The Japan Biological Informatics Consortium and the National Institute of Advanced Industrial Science and Technology established one of the world's largest natural-substance libraries, with natural-substance libraries provided by pharmaceutical companies and other entities in a NEDO project, "Development of Basic Technologies for Biological System Control by Using Chemical Compounds"
April 2011	The Association was founded to facilitate industrial applications, such as drug development, by constantly keeping and controlling one of the world's largest natural-substance libraries to promote the mutual use of the libraries
	Contracted a NEDO project, "Development of Technology for Stable Production of Useful Natural Compounds (2011-2012)"
August 2013	Contracted a METI project, "Development of Innovative Drug Discovery Platform for Personalized Medicine (Technology for Creating Innovative Drugs Utilizing Natural Compounds and Information Technology)"
April 2014	In accordance with the basic plan revision in January 2014, the project title was changed to "Development of Drug Discovery Technology for New-Generation Treatments and Diagnoses (Technology for Creating Innovative Drugs Utilizing Natural Compounds and Information Technology)"
January 2015	Contracted duties related to use of natural-substance libraries in the "drug discovery support network" from the National Institute of Biomedical Innovation
April 2015	Upon the founding of the Japan Agency for Medical Research and Development (AMED), contracted "Development of Drug Discovery Technology for New-Generation Treatments and Diagnoses (Technology for Creating Innovative Drugs Utilizing Natural Compounds and Information Technology)" as an AMED project
April 2018	Contracted from AMED to undertake its project, "Development of Drug Discovery Technology for Next-Generation Treatments and Diagnoses (Development of Technology for Creating Innovative Intermediate-Molecular Drugs)"

Acceptable contract operations (feel free to ask any questions)

- **Sequence analysis**
- **Contract preparation of BAC libraries**

  Developing a BAC library from bacterial producers of useful natural compounds on a contract basis

  Because cloning is achieved using BAC, long genes can be cloned as they are.
- **Production of secondary microbial metabolites by heterologous expression**

  1. Microbial genome decoding
  2. Preparation of genome libraries using BAC vectors
  3. Cloning of desired biosynthesis gene
  4. Production by heterologous expression in clean host strains
  5. Core modification by biosynthesis gene modification
- **Module editing**

  Preparation of analogues of useful natural compounds

Technology Research Association for Next-generation Natural Products Chemistry

TIME 24 Building 10F, 2-4-32, Aomi, Koto-ku, Tokyo, 135-8073 Japan
TEL: 03-5531-8554 FAX: 03-5531-1560
Homepage: http://www.natprodchem.jp/
E-mail: natprodchem@jbic.or.jp

Issued in March 2020,
Ver. 1.1