

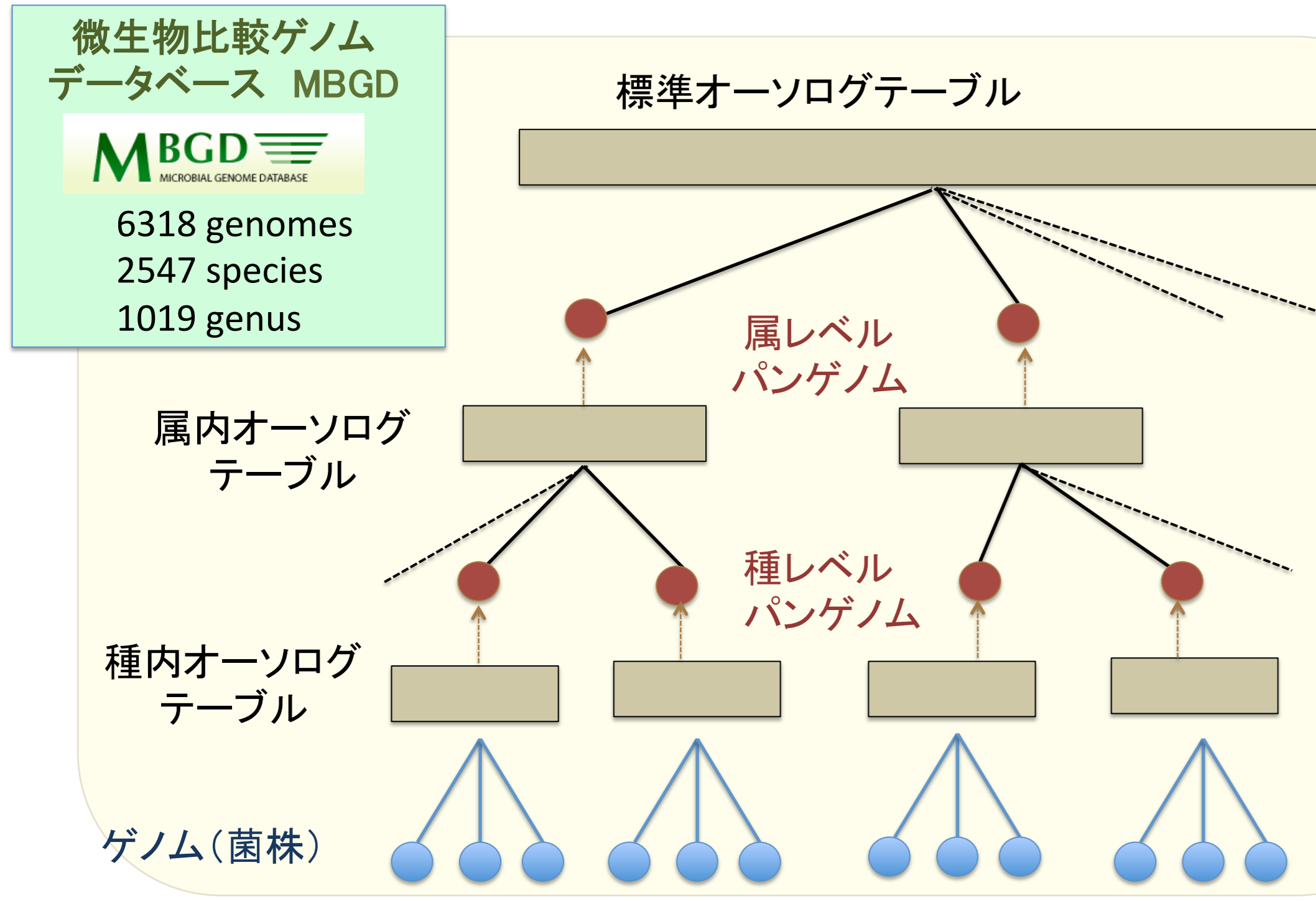


# 微生物比較ゲノムデータベースMBGD

内山郁夫<sup>1</sup>、三原基広<sup>2</sup>、西出浩世<sup>1</sup>、千葉啓和<sup>3</sup>、高柳正彦<sup>4</sup>、高見英人<sup>5</sup>  
(1.基生研, 2.ダイナコム, 3. DBCLS, 4.ウェブブレイン, 5. JAMSTEC)

## 遠縁・近縁オーソログテーブルの階層的統合

MBGDでは、全系統群にまたがる標準オーソログテーブルに加えて、各系統群に特異的なオーソログテーブルを作成している。最新版では種内比較、属内比較、属間比較を段階的に行い、下位のオーソログ解析から得られる遺伝子レパートリー全体(パンゲノム)を、上位の比較の入力に用いるアプローチで階層的に結合することにより、全遺伝子レパートリーを網羅するオーソログテーブルの構築を行っている。



代表配列ベースからパンゲノムベースへ変更したときの標準テーブルサイズの変化

クラスター数  
491,920 → 768,073  
配列数  
3,735,085 → 4,640,598

トータルの配列数  
22,521,946

## オーソログテーブルブラウザ

オーソログテーブルブラウザでは、従来通りに全体をカバーする標準オーソログテーブルと系統特異的なオーソログテーブルを切り替えて表示できるが、階層関係にある種・属レベルのオーソロググループ間では、相互にリンクによって辿ることができる。

### サマリビュー

## キーワード検索

オーソロググループ、遺伝子、生物種に対する新しいキーワード検索機能として、ElasticsearchおよびjQueryを用いた高速・高機能な検索を実現した。

## MBGDトップページ

<http://mbgd.genome.ad.jp/>

## 配列プロファイル検索

MBGD標準オーソログテーブルのオーソロググループ(3メンバー以上)のマルチプルアライメントに基づいて、プロファイルライブラリを作成しており、それを用いてユーザクエリ配列に対するHMMERまたはMMseqsによるプロファイル検索が行える。特にMMseqsを用いると、ゲノム規模の大量のクエリ配列に対しても、高速に(数分程度)検索を行うことができる。

## 生物種・系統群検索

## クエリ配列入力

## クラスター検索/遺伝子検索

## 配列検索結果

NAME	HIT	E-value	Score	Identity	Region	Description
lklLokArch_00020	634	1.06E-101	342	0.383		Preprotein translocase subunit SecY
lklLokArch_00030	7655	1.00E-17	76	0.468		50S ribosomal protein L12
lklLokArch_00040	6612	5.73E-55	196	0.472		Acidic ribosomal protein P0
lklLokArch_00050	633	3.70E-13	71	0.286		50S ribosomal protein L1
lklLokArch_00060	609	5.53E-25	105	0.411		50S ribosomal protein L11
lklLokArch_00070	9675	1.06E-40	154	0.463		PAS/PAC sensor signal transduction histidine kinase
lklLokArch_00070	4016	2.57E-15	75	0.640		Sensor histidine kinase/response regulator
lklLokArch_00070	1446	1.43E-13	71	0.232		Sensor signal transduction histidine kinase
lklLokArch_00080	3181	1.07E-15	72	0.465		CheY chemotaxis protein, response regulator
lklLokArch_00100	287	1.47E-08	54	0.497		AraP family transcriptional regulator

## MAPLEとの統合によるゲノム機能解析アプリケーション

ユーザクエリゲノム配列解析機能をさらに強化するため、現在 JAMSTEC 高見グループにより開発されている MAPLE (Metabolic And Physiological potential Evaluator) との統合を進めている。これにより、ユーザゲノムに存在する機能的特徴を機能モジュール(KEGGモジュール)の充足率から推定し、その結果をMBGDの比較ゲノム解析機能を使って、複数の他ゲノムと詳細に比較することが可能になる。

## 系統プロファイル検索

系統プロファイル(ゲノムごとの遺伝子の有無のパターン)検索のためのインターフェイスも新たに開発している。これにより、系統群、表現形質のほか、MicrobeDB.jpに登録された、メタゲノム解析から得られた各環境における存在パターンなどをクエリパターンとして指定し、関連するオーソロググループを検索することが可能になる。

## ユーザゲノム登録(MyMBGD)

## MMseqs 検索

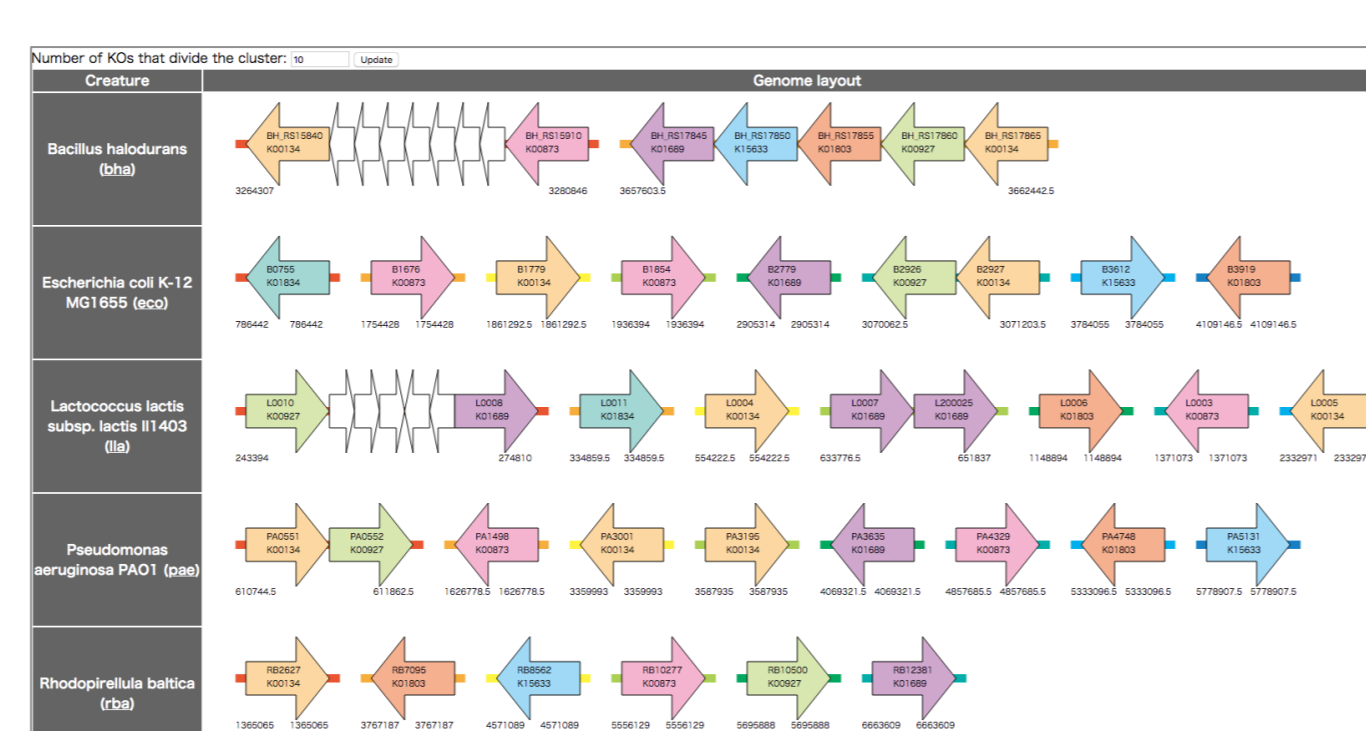
## MBGD-KEGG Orthology クロスリファレンス

## MAPLE解析

## KEGGモジュールリスト

## KEGGモジュール構成比較画面

## モジュール構成遺伝子のゲノム上の並び比較画面



## パターン指定画面 運動性(Motility)のあるなしをクエリとして検索する例