

# 平成 26 年度 研究開発実施報告書

ライフサイエンスデータベース統合推進事業「統合化推進プログラム」研究代表者

[黒川 顕]

[東京工業大学 地球生命研究所・教授]

[ゲノム・メタゲノム情報統合による微生物DBの超高度化推進]

## §1. 研究実施体制

### (1) 東工大グループ(研究機関別)

- ① 研究代表者:黒川 顕 (東京工業大学地球生命研究所、教授)
- ② 研究項目:メタゲノムおよび微生物 DB システムの高度化と統合解析システムの開発

### (2) 遺伝研グループ(研究機関別)

- ① 研究代表者:中村 保一 (国立遺伝学研究所生命情報研究センター、教授)
- ② 研究項目:微生物 DB の高度化と配列解析支援系の開発と提供

### (3) 基生研グループ(研究機関別)

- ① 主たる共同研究者:内山 郁夫 (基礎生物学研究所・ゲノム情報研究室、助教)
- ② 研究項目:比較ゲノム解析に立脚した微生物ゲノム情報の統合化

## § 2. 研究実施内容

### 1. メタゲノムおよび微生物 DB システムの高度化と統合解析システムの開発 (東工大 G)

#### 1-1. メタゲノム DB の整備

##### 1) 各種オミックスデータへの対応

SRA に収録されている RNA-Seq データについて、大腸菌をテストデータとして Reference ゲノムにリードデータをマッピングした結果を基に、ゲノム中の遺伝子ごとの発現量のデータを RDF 化するための RDF モデルをデザインした。

##### 2) 各種オントロジー、ボキャブラリの開発

拡大するメタゲノムデータに対応してより多様な環境情報を記述できるように、SRA・IMG/M・MG-RAST の 3 つの主要なメタゲノム DB から取得したメタゲノムデータ中の生息環境情報から新規語彙を抽出し、それらを微生物の生息環境オントロジーである Metagenome/Microbes environmental ontology (MEO) に追加すると共に MEO の階層構造を変更し、MEO version 0.8 を構築した。MEO version 0.8 を用いて、SRA・IMG/M・MG-RAST のメタゲノムデータの生息環境情報を、MEO を用いて再アノテーションを行った上 RDF 化した。

#### 1-2. 解析 Stanza の開発

H26 年度は、高度な解析で得られた結果を表示できる各種 Stanza を開発し、これまで開発したもの、TogoStanza により開発されたものと併せて合計 160 を超える Stanza を開発した。特筆すべきものとして、入力したメタ 16S サンプル群の系統組成データに対して PCoA 解析を行い 2 次元の刺激布置を描く Stanza や、サンプル間距離を Bray-Curtis distance など で計算し階層的クラスタリングを行う Stanza を開発した。これら解析 Stanza を MicrobeDB.jp に実装することにより、メタ 16S 解析の基本的な一連の解析が MicrobeDB.jp 上で可能となる。さらに、各メタ 16S・メタゲノムデータの系統組成と MEO による生息環境のアノテーション情報をレファレンスとして用いて、ユーザが入力したメタ 16S・メタゲノムサンプルと最も類似した系統組成を持つサンプルを特定し、そのサンプルの環境情報をユーザに提案し、細菌群集の系統組成を用いて類似した環境を推定するアプリケーションも開発した。

#### 1-3. 真菌類および藻類データの整備

DBCLS と共同で、ゲノムの RDF モデルを真核生物および原核生物に共通して使用出来るように再設計した。真核生物のオミックスデータについては、代表的な真核藻類であり研究が進んでいる *Chlamydomonas reinhardtii* をモデル生物として、JGI Phytozome からゲノムデータ、SRA から RNA-Seq データ (約 3.5TB) をそれぞれ取得し、ゲノムに全ての RNA-Seq のリードデータをマッピングして実験ごとに各遺伝子の発現量を取得した。

#### 1-4. DB 更新自動化システムの開発

1) データの収集およびクオリティコントロール (QC)、更新の自動化など持続可能なシステムの構築およびデータ共有、公開におけるアクセスレベルの制御システムの構築

ユーザが所持するゲノム・メタゲノムデータを受け入れるために、微生物ゲノムのオートアノテーションパイプライン MiGAP およびメタゲノムの解析パイプライン MeGAP と、ゲノム・メタゲノムデータ

およびユーザが入力したそれらのデータに付随するメタデータを RDF に変換し MicrobeDB.jp に転送する機能を持つ Web アプリケーションである GenomeRefine 間の、データの自動受け渡し機能、および GenomeRefine-MicrobeDB.jp 間での OpenID を用いたユーザ認証機能およびユーザを束ねるグループ管理機能を開発した(遺伝研 G と共同)。さらに、DRA/ERA/SRA のメタ 16S とメタゲノムサンプルの fastq ファイル 143,378 データについて、DDBJ スパコン上で新解析パイプラインを実装し、サンプルごとに系統組成を推定する計算を開始した。

## 2) ユーザビリティの向上

ユーザが入力した検索キーワードから、MicrobeDB.jp の中にあるデータ中で完全マッチおよび部分マッチするキーワードを遺伝子・系統・環境・表現型等のいくつかのカテゴリごとに列挙し、よりユーザが検索したい情報にたどり着きやすくする検索システムのテスト版を開発した。

## 2. 微生物 DB の高度化と配列解析支援系の開発と提供(遺伝研 G)

各種オミックスデータの整備と藻類データを整備することにより微生物 DB の高度化をさらに進めるとともに、配列解析支援系としてゲノム・メタゲノム配列解析パイプラインである MiGAP および MeGAP との MicrobDB.jp の連携を強化し、アクセスレベルの制限システムの構築を実施することで微生物ゲノム・メタゲノム配列の価値を高め、生物学者のデータベースへの情報集積ならびにその高度な利用の促進を図ることを研究開発の目的とし、開発を実施した。

### 2-1. 各種オミックスデータの整備

これまでの微生物ゲノム情報 DB に不足していた各種オミックスデータを統合的に整理し提供するための枠組みの整備を目標に、MicrobeDB.jp にオミックスデータを対象としたシステム整備と速やかな収集およびそのセマンティックリソース化をすすめた。H26 年度は本システム整備に着手すると共に、コミュニティとの連携をもとに主要モデル微生物のゲノム情報およびメタデータに関する INSDC リソースを中心に NCBI genome reports, assembly reports, BioSample, BioProject のデータ収集を実施した。また、収集したデータについては、これまでに開発したゲノム情報 RDF 等とのデータ統合し、SPARQL による検索性を高めるために RDF データモデルの設計および開発を行なった。

### 2-2. 藻類データの整備

単細胞藻類のゲノム情報ならびにオミックスデータを MicrobeDB.jp 上に収集・整備するための準備段階として、利用可能な藻類データとそのデータモデルについて DBCLS と PGDBj と連携しつつ、オントロジーの共用化を中心に情報収集と整理を行った。

### 2-3. MiGAP および MeGAP との連携強化

開発したウェブサービス GenomeRefine (<http://genome.annotation.jp/genomerefine/>) をセマンティックウェブ技術およびリソースを用いて機能強化し MiGAP との連携を実現することで、MiGAP の解析出力を MicrobeDB.jp ならびに INSDC 公共 DB 受入システムへのスムーズな登録を支援するシステムとして、MeGAP 用メタゲノム解析メタデータ受け入れのための INSDC データベース登録情報に対応したメタゲノム解析用 SRA メタデータ入力項目を定義し、拡充整備した。

H26 年度は GenomeRefine のセマンティックウェブ技術による強化に着手し DDBJ の提供する RDF 等、外部 DB との共用可能なリソースとの連動を整備した(東工大 G と共同)。

#### 2-4. アクセスレベルの制限システムの構築

生物学者の利用と登録を促進し、また、論文公開後には速やかに配列ならびに解析情報を公開し利用に供することを目標に MicrobeDB.jp 上でアクセスコントロール環境を提供し、制限条件下での配列登録や解析をサポートする開発を実施した。これまでに実施した OpenID の拡張によりグループの概念を付与した認証システムをさらに拡充し、MicrobeDB.jp 全域に適用する開発を実施することで、グループ単位でのアクセスレベルコントロールを実現し、配列解析者の利便性を増すと同時に、配列公開の促進を図った。H26 年度は MicrobeDB.jp 上の各種認証機能に対して拡張 OpenID 認証システムの適用を行うとともに、API 整備とドキュメント整備を実施し、すでに統合 DB プロジェクト内で利用されている OpenID 認証機能を本拡張による認証に置き換えるために必要な基盤整備を実施した。

### 3. 比較ゲノム解析に立脚した微生物ゲノム情報の統合化(基生研 G)

微生物比較ゲノムデータベース MBGD を拡張して、単細胞真核生物を含む多様な微生物の比較解析を行えるようにする。特にドラフト状態で公開されたゲノムについても取り込んで、比較解析に利用できるようにするとともに、複雑なドメイン構成を持つオーソログ分類の分類精度を改善する。また、オーソログ情報に基づいて各種データを統合するため、オーソログ関連データの RDF 化と検索インターフェイスの開発を進める。H26 年度は、これまでに開発してきたオーソログ分類の精密化手法 DomRefine、ならびにオーソログ情報を RDF 化するためのオントロジー OrthO と検索のための SPARQL 検索サイトなどの研究成果を論文発表し、MBGD のサービスに取り込んで公開した。また、ドラフトゲノムを取り込んで解析する仕組みについても MBGD 上に実装し、利用者インターフェイスを介して利用できるよう設定して公開した。

#### 3-1. DomRefine を用いた標準オーソログテーブルの精密化

MBGD で用いている DomClust はドメイン単位でオーソログ分類を行っているが、ペアワイズアライメントに基づいてドメイン分割を行っているために、不正確に分割されるケースがあった。そこで、マルチプルアライメントを用いてドメイン分割の精度を改善する DomRefine プログラムを開発した<sup>1)</sup>。DomRefine の第 1 バージョンの開発とその評価は前年度までに概ね完了していたが、若干の高速化を行った上でこれを MBGD の更新プロセスに組み込んで、MBGD の標準オーソログテーブルの精密化を行うようにした<sup>2)</sup>。この過程で計算されるマルチプルアライメントや系統樹の情報もストアしておき、取得できるようにした。

#### 3-2. ドラフトゲノムの取り込みとオーソログ解析への活用

真核生物・原核生物含め、近年公開されるゲノムの多くがドラフト状態で公開されるようになってきているため、それらを効果的に活用するためのデータ処理方式を構築した。MBGD では、属(genus)当たり 1 つを代表にとった生物種セットで作成したオーソログテーブルをデフォルト(標準オーソログテーブル)として使っている。これに合わせて、ドラフトゲノムのうち、質が比較的高く、標

準オーソログテーブルには含まれない属に属するものは MBGD に追加し、標準オーソログテーブルを拡張した **draft-plus** テーブルとして利用できるようにした。それ以外のドラフトゲノムについては、データベースに取り込むが事前のホモロジー計算の対象にはせず、利用者が MyMBGD 機能を通じて動的に解析に加えることができるようにした。この方針は、前年度までに確立していたが、利用者インターフェイスが整わず、公開までには至っていなかった。H26 年度に MyMBGD インターフェイスの全面的な改定を行い、これらの機能を利用できるようにして公開した<sup>12)</sup>。一方、MBGD では従来、主に NCBI RefSeq の genomes セクションから完全ゲノムデータを取り込んできたが、NCBI では完全ゲノムとドラフトゲノム、原核生物と真核生物とを合わせて、すべて Assembly report を介して公開する体制に移行してきており、それに合わせてデータの取得・構築プロセスの見直しを行い、全面的な修正を行った。

### 3-3. オーソログ情報の RDF 化

オーソログ情報は、様々な生物種において蓄積した遺伝子機能や表現型などの知識を相互に結びつけるための中枢として機能しうる。セマンティックウェブ技術を用いて、MicrobeDB.jp の他のデータベースと連携して MBGD のオーソログ情報を検索できるようにすることを目指して、オーソログ情報の RDF 化を行った。この目的のため、オーソログの概念を定義したオントロジー Ortholog Ontology (OrthO)を開発し、これを用いて MBGD のオーソログ情報を RDF 化した<sup>13)</sup>。また、SPARQL 検索のためのポータルサイトを立ち上げて公開し、UniProt や Taxonomy データベースなどと統合的に検索できるようにした<sup>13)</sup>。

## §3. 成果発表等

### (3-1) 原著論文発表

- ① 発行済論文数(国内(和文) 0 件、国際(欧文) 13 件)
  - ② 未発行論文数(“accepted”、“in press”等)(国内(和文) 0 件、国際(欧文)0 件)
  - ③ 論文詳細情報
1. Yano M, Mori H, Akiyama Y, Yamada T, Kurokawa K. CLAST: CUDA implemented large-scale alignment search tool. *BMC Bioinformatics*, 15:406, 2014. (DOI: 10.1186/s12859-014-0406-y)
  2. Wang T, Mori H, Zhang C, Kurokawa K, Xing X, Yamada T. DomSign: a top-down annotation pipeline to enlarge enzyme space in the protein universe. *BMC Bioinformatics*, 16:96, 2015. (DOI: 10.1186/s12859-015-0499-y)
  3. Kojima H, Ogura Y, Yamamoto N, Togashi T, Mori H, Watanabe T, Nemoto F, Kurokawa K, Hayashi T, Fukui M. Ecophysiology of *Thioploca ingrica* as revealed by the complete genome sequence supplemented with proteomic evidence. *ISME J.*, 9:1166–1176, 2015. (DOI: 10.1038/ismej.2014.209)
  4. Aoki-Kinoshita KF, Kinjo AR, Morita M, Igarashi Y, Chen Y, Shigemoto Y, Fujisawa T, Akune Y, Katoda T, Kokubu A, Mori T, Nakao M, Okamoto S, Katayama T, Ogishima S. Implementation of linked data in the life sciences at BioHackathon

2011. *Journal of Biomedical Semantics* 2015, 6:3. (DOI: 10.1186/2041-1480-6-3)
5. Kanesaki, Y Masutani H, Sakanaka M, Shiwa Y, Fujisawa T, Nakamura Y, Yokota A, Fukiya S, Suzuki T, Yoshikawa H. Complete genome sequence of *Bifidobacterium longum* 105-A, a strain with high transformation efficiency. *Genome Announc.* 2014, 2(6):e01311-14. (DOI: 10.1128/genomeA.01311-14)
  6. Kodama Y, Mashima J, Kosuge T, Katayama T, Fujisawa T, Kaminuma E, Ogasawara O, Okubo K, Takagi T, Nakamura Y. The DDBJ Japanese Genotype-phenotype Archive for genetic and phenotypic human data. *Nucleic Acids Res.* 2015, 43:D18-22. (DOI: 10.1093/nar/gku1120)
  7. Hiraide Y, Oshima K, Fujisawa T, Uesaka K, Hirose Y, Tsujimoto R, Yamamoto H, Okamoto S, Nakamura Y, Terauchi K, Omata T, Ihara K, Hattori M, Fujita Y. Loss of cytochrome cM stimulates cyanobacterial heterotrophic growth in the dark. *Plant Cell Physiol.* 2015, 56(2):334-45. (DOI: 10.1093/pcp/pcu165)
  8. Tanizawa Y, Fujisawa T, Mochizuki T, Kaminuma E, Suzuki Y, Nakamura Y, Tohno M. Draft genome sequence of *Weissella oryzae* SG25T, isolated from fermented rice grains. *Genome Announc.* 2014, 2(4):e00667-14. (DOI:10.1128/genomeA.00667-14)
  9. Tanizawa Y, Fujisawa T, Mochizuki T, Kaminuma E, Nakamura Y, Tohno M. Draft genome sequence of *Lactobacillus oryzae* strain SG293T. *Genome Announc.* 2014, 2(4):e00861-14. (DOI: 10.1128/genomeA.00861-14)
  10. Hori K, Maruyama F, Fujisawa T, Togashi T, Yamamoto N, Seo M, Sato S, Yamada T, Mori H, Tajima N, Moriyama T, Ikeuchi M, Watanabe M, Wada H, Kobayashi K, Saito M, Masuda T, Sasaki-Sekimoto Y, Mashiguchi K, Awai K, Shimojima M, Masuda S, Iwai M, Nobusawa T, Narise T, Kondo S, Saito H, Sato R, Murakawa M, Ihara Y, Oshima-Yamada Y, Ohtaka K, Satoh M, Sonobe K, Ishii M, Ohtani R, Kanamori-Sato M, Honoki R, Miyazaki D, Mochizuki H, Umetsu J, Higashi K, Shibata D, Kamiya Y, Sato N, Nakamura Y, Tabata S, Ida S, Kurokawa K, Ohta H. *Klebsormidium flaccidum* genome reveals primary factors for plant terrestrial adaptation. *Nat Commun.* 2014 May 28;5:3978. (DOI: 10.1038/ncomms4978)
  11. Chiba H, Uchiyama I. “Improvement of domain-level ortholog clustering by optimizing domain-specific sum-of-pairs score”, *BMC Bioinformatics*, 15, 148, 2014. (DOI: 10.1186/1471-2105-15-148)
  12. Uchiyama I, Mihara M, Nishide H, Chiba H. “MBGD update 2015: microbial genome database for flexible ortholog analysis utilizing a diverse set of genomic data”, *Nucleic Acids Res.*, 43, D270-276, 2015. (DOI: 10.1093/nar/gku1152)
  - Chiba H, Nishide H, Uchiyama I. “Construction of an ortholog database using the Semantic Web technology for integrative analysis of genomic data”, *PLoS ONE*, 10, e0122802, 2015. (DOI: 10.1371/journal.pone.0122802)

(3-2) データベースおよびウェブツール等の構築と公開

別紙1参照。

### (3-3) その他の著作物(総説、書籍など)

#### ●詳細情報

1. 山本希, 森宙史, 山田拓司, 黒川顕, “メタゲノミクスの現状と未来”, 生命のビッグデータ利用の最前線, シーエムシー出版, pp48-57, 2014年.
2. 森宙史, 黒川顕, “MicrobeDB.jp”, 今日から使えるデータベース・ウェブツール, 羊土社, pp76-77, 32(20), 2014年.

### (3-4) 国際学会および国内学会発表

#### ① 招待講演 (国内 9件、国際 2件)

##### 〈国内〉

1. 森宙史, 古代ゲノムと微生物統合データベース, 北海道大学低温科学研究所研究集会「低温環境への適応:ゲノム進化と環境変動指標」, 北海道大学, 2014年7月22-23日.
2. 黒川顕, 統合データベースを活用した冥王代類似環境微生物のゲノム情報解析, 大阪, 日本進化学会第16回大阪大会, 2014年8月23日.
3. 森宙史, 藤澤貴智, 大山彰, 菅原秀明, 中村保一, 黒川顕, MiGAPを用いた微生物ゲノムオートアノテーションとMicrobeDB.jp連携, 第8回日本ゲノム微生物学会若手の会, ろうきん研修所富士センター, 2014年9月28-29日.
4. 内山郁夫, 微生物比較ゲノムデータベース MBGD, 第3回生命医薬情報学連合大会・オープンサイエンスアワード2014, 仙台国際センター, 2014年10月2-4日.
5. 黒川顕, 微生物統合データベース「MicrobeDB.jp」, 環境微生物系学会合同大会2014, 浜松, 2014年10月22日.
6. 森宙史, 細菌群集解析のためのメタ16S・メタゲノム解析パイプラインの開発, Illumina BaseSpace Worldwide Developers Conference in Japan, 東京大学柏の葉キャンパス駅前サテライト, 2014年10月29-30日.
7. 森宙史, メタゲノム解析と微生物統合データベース, 第30回DDBJing講習会in東京, 独立行政法人科学技術振興機構東京本部, 2014年12月18日.
8. 黒川顕, 微生物統合データベース「MicrobeDB.jp」, 第88回日本細菌学会総会, 長良川国際会議場, 2015年3月26日.
9. 内山郁夫, 属、科など上位分類群に保存されたコア遺伝子の定義と分類への応用, 第88回日本細菌学会総会, 長良川国際会議場, 2015年3月26-28日.

##### 〈国際〉

1. Ken Kurokawa, “Small Bugs, Big Data”: Developing an integrated database for microbes with semantic web technologies, ICSTI2014 General Assembly & Annual Conference in Tokyo, Tokyo, 2014年10月20日.
2. Hirokazu Chiba, Hiroyo Nishide, Ikuo Uchiyama, Construction of ortholog database



using the Semantic Web technology BioHackathon2014 Symposium, 東北大学, 2014年11月9日.

② 口頭講演 (国内 3件、国際 0件)

〈国内〉

1. 藤澤貴智, 中村保一, ゲノムアノテーション支援サービス「Genome Refine」の開発と現状, ラン藻ゲノム研究交流会 2014, 東京, 2014年7月19日.
2. 内山郁夫, 三原基広, 西出浩世, 千葉啓和, 大規模ゲノムデータの活用に向けた微生物比較ゲノムデータベース MBGD の改良, 第9回日本ゲノム微生物学会年会, 神戸大学, 2015年3月6-8日.
3. 森宙史, 内山郁夫, 菅原秀明, 中村保一, 黒川顕, MicrobeDB.jp プロジェクトチーム, 微生物統合データベース MicrobeDB.jp の超高度化, 第9回日本ゲノム微生物学会年会, 神戸大学, 2015年3月6-8日.

③ ポスター発表 (国内 3件、国際 1件)

〈国内〉

1. 藤澤貴智, 森宙史, 岡本忍, 山本泰智, 片山俊明, 川島秀一, 谷澤靖洋, 神沼英里, 大山彰, 菅原秀明, 内山郁夫, 黒川顕, 中村保一, Genome Refine: 自動ゲノムアノテーション統合環境を提供するウェブサービス, 第9回日本ゲノム微生物学会年会, 神戸大学, 2015年3月6-8日.
2. 鈴木真也, 山本希, 森宙史, 黒川顕, 微生物統合データベースの超高度化推進のための微生物の生息環境オントロジーと解析アプリケーションの開発, 第9回日本ゲノム微生物学会年会, 神戸大学, 2015年3月6-8日.
3. Hirokazu Chiba, Hiroyo Nishide, Ikuo Uchiyama, Construction of ortholog database using the Semantic Web technology, JSBi2014・第3回生命医薬情報学連合大会, 仙台国際センター, 2014年10月2-4日.

〈国際〉

1. Hiroshi Mori, Nozomi Yamamoto, Takuji Yamada, Ken Kurokawa, MicrobeDB.jp: Integration of a microorganism database based on genomic and metagenomic information, the 3rd ELSI International Symposium, Kuramae Hall, Tokyo, Japan, January 13-15, 2015.

(3-6) 知的財産権の出願

- ① 当該年度の特許出願件数(国内 0件、海外 0件)
- ② 研究開始から当該年度末までの累積特許出願件数(国内 0件、海外 0件)

③ その他の知的財産権

無し。

(3-7) 受賞・報道等

無し。

§4. 研究開発期間中に主催した活動(主催したワークショップ等)

年月日	名称	場所	参加人数	概要
2014年7月5日	チーム内第1回全体会議(非公開)	東工大田町キャンパス	10人	研究進捗報告のための会議
2015年2月21日	チーム内第2回全体会議(非公開)	東工大大岡山キャンパス	8人	研究進捗報告のための会議

以上

別紙1 既公開のデータベース・ウェブツール等

No.	研究開発課題名 データベース、ウェブツールの名称【公開】	概要【公開】 (150字程度)	URL【公開】	公開開始日	関連論文 (論文リストに記載があれば、その番号でも可)	備考
01	ゲノム・メタゲノム情報統合による微生物DBの超高度化推進 MicrobeDB.jp	本DBは、微生物に関する多種多様な情報を遺伝子・系統・環境の3つの軸に沿って整理統合し、セマンティックWeb技術を利用して単一の検索ウィンドウからそれらの情報を検索可能な統合DBです。	http://microbedb.jp/	2011/12/12		
02	ゲノム・メタゲノム情報統合による微生物DBの超高度化推進 CyanoBase (MicrobeBase)	シアノバクテリアに代表される酸素発生型光合成細菌とその関連バクテリアのゲノム情報を集積したデータベースです。	http://genome.microbedb.jp/cyanobase	1995/12/1	Takatomo Fujisawa, Shinobu Okamoto, Toshiaki Katayama, Mitsuteru Nakao, Hidehisa Yoshimura, Hiromi Kajiya-Kanegae, Sumiko Yamamoto, Chiyoiko Yano, Yuka Yanaka, Hiroko Maita, Takakazu Kaneko, Satoshi Tabata, Yasukazu Nakamura, "Nucleic Acids Res.", 42:D666-70, 2014 (DOI: 10.1093/nar/gkt1145)	
03	ゲノム・メタゲノム情報統合による微生物DBの超高度化推進 GTPS	本DBでは微生物ゲノムの最新情報を基にした再アノテーション情報を提供しています。	http://gtps.ddbj.nig.ac.jp/gtps_data/	2009/3/1	Takehide Kosuge, Takashi Abe, Toshihisa Okido, Naoto Tanaka, Masaki Hirahata, Yutaka Maruyama, Jun Mashima, Aki Tomiki, Motoyoshi Kurokawa, Ryutaro Himeno, Satoshi Fukuchi, Satoru Miyazaki, Takashi Gojobori, Yoshio Tateno and Hideaki Sugawara, "Exploration and Grading of Possible Genes from 183 Bacterial Strains by a Common Protocol to Identification of New Genes: Gene Trek in Prokaryote Space (GTPS)", DNA Res. 13(6):245-254, 2006 (DOI:10.1093/dnares/ds1014)	
04	ゲノム・メタゲノム情報統合による微生物DBの超高度化推進 GTPS/RDF	本DBはセマンティックWeb技術によって、GTPSと微生物統合データベースにおける他の要素データベースとの連携を可能としたGTPS2011のRDF化版です。	http://gtps.ddbj.nig.ac.jp/rdf/	2012/4/26		
05	ゲノム・メタゲノム情報統合による微生物DBの超高度化推進 Metagenome.jp (Human Meta BodyMap)	世界中で公開されているヒトメタゲノム解析で得られた配列情報を、独自に注意深く再アノテーションしたメタデータに基づき参照できるようにしたDBです。メタデータ検索機能や配列同源性検索ツールBody-BLASTを用いたメタデータ検索機能も提供しています。	http://www.metagenomics.jp/mg/	2012/3/25		
06	ゲノム・メタゲノム情報統合による微生物DBの超高度化推進 MBGD	オーソログ解析に基づいて微生物ゲノムの比較解析を行うためのデータベースです。公開されたゲノム全体を含む標準オーソログテーブルに基づいて、各オーソロググループの系統プロファイルの比較などを行えるほか、動的なオーソログ解析機能によって、利用者自身が持つゲノム配列も含めて、興味のある生物種セットを対象とした比較を行うこともできます。	http://mbgd.genome.ad.jp/	1997/7/1	Ikuo Uchiyama, Motohiro Mihara, Hiroyo Nishide, and Hirokazu Chiba, "MBGD update 2015: microbial genome database for flexible ortholog analysis utilizing a diverse set of genomic data", Nucleic Acids Res., 43, D270-276, 2015. (DOI: 10.1093/nar/gku1152)	
07	ゲノム・メタゲノム情報統合による微生物DBの超高度化推進 MEO	本DBは、微生物の生息環境に関するメタデータを記述し整理するためのオントロジーであるMEO (Metagenome/Microbes Environmental Ontology)のOWLファイルと、公共のメタゲノムデータのメタデータをMEO相手にマッピングしてサンプルごとに整理した結果のRDFファイルの2つから構成されたDBです。	http://mdb.bio.titech.ac.jp/meo	2012/1/31		
08	ゲノム・メタゲノム情報統合による微生物DBの超高度化推進 RhizoBase (MicrobeBase)	窒素固定植物共生細菌である根粒菌(共生関連領域のみの部分配列を含む)ゲノム情報を集積したデータベースです。	http://genome.microbedb.jp/cyanobase	2001/1/1	Takatomo Fujisawa, Shinobu Okamoto, Toshiaki Katayama, Mitsuteru Nakao, Hidehisa Yoshimura, Hiromi Kajiya-Kanegae, Sumiko Yamamoto, Chiyoiko Yano, Yuka Yanaka, Hiroko Maita, Takakazu Kaneko, Satoshi Tabata, Yasukazu Nakamura, "CyanoBase and RhizoBase: databases of manually curated annotations for cyanobacterial and rhizobial genomes", Nucleic Acids Res., 42:D666-70, 2014 (DOI: 10.1093/nar/gkt1145)	
09	ゲノム・メタゲノム情報統合による微生物DBの超高度化推進 TogoAnnotation	ソーシャルブックマークによる文献情報集積プラットフォームです。	http://togo.annotation.jp/	2007/6/1	Takatomo Fujisawa, Shinobu Okamoto, Toshiaki Katayama, Mitsuteru Nakao, Hidehisa Yoshimura, Hiromi Kajiya-Kanegae, Sumiko Yamamoto, Chiyoiko Yano, Yuka Yanaka, Hiroko Maita, Takakazu Kaneko, Satoshi Tabata, Yasukazu Nakamura, "CyanoBase and RhizoBase: databases of manually curated annotations for cyanobacterial and rhizobial genomes", Nucleic Acids Res., 42:D666-70, 2014 (DOI: 10.1093/nar/gkt1145)	
10	ゲノム・メタゲノム情報統合による微生物DBの超高度化推進 PDO-CSSO	本DBは、微生物がヒトに引き起こす感染症の病名を、ヒトの体の組織ごとにまとめて記述・整理したオントロジーであるPDO (Pathogenic Disease Ontology)のOWLファイルと、感染症が引き起こす症状についてのオントロジーであるCSSO (Clinical Signs and Symptoms Ontology)のOWLファイル、さらに公共のゲノムデータのメタデータをPDO-CSSO相手にマッピングし、宿主となる生物の系統名と、その感染症に関連する細菌名とを関連づけてゲノムごとに整理した結果のRDFファイルから構成されたDBです。	http://mdb.bio.titech.ac.jp/pdo http://mdb.bio.titech.ac.jp/csso	2013/5/31		
11	ゲノム・メタゲノム情報統合による微生物DBの超高度化推進 MCCV	本DBは、JCMやNBRC等の、菌株保存機関に蓄積されている菌株情報を記述するための統制語彙であるMCCV (Microbial Culture Collection Vocabulary)のOWLファイルを提供するDBです。DBCLSと共同で開発しております。	http://bioportal.bioontology.org/ontologies/MCCV	2013/1/31		
12	ゲノム・メタゲノム情報統合による微生物DBの超高度化推進 Genome Refine	微生物ゲノムアノテーションの品質向上および解析支援のためのウェブサービスです。ユーザ認証機能、メタデータ入力機能、解析パイプラインMiGAP、MeGAPの解析結果の登録機能、RDF形式変換機能を持ち、MicrobeDB.jpで参照可能です。	http://genome.annotation.jp/genomerefine/	2014/7/1		
13	ゲノム・メタゲノム情報統合による微生物DBの超高度化推進 MBGD SPARQL Search	MBGDのRDF版をSPARQLで検索するためのウェブサイトです。SPARQLによる検索例や、RDFアーカイブ、Ortholog Ontology (Ortho)のドキュメントなども含んでいます。	http://mbgd.genome.ad.jp/sparql	2014/10/1	Hirokazu Chiba, Hiroyo Nishide, and Ikuo Uchiyama, "Construction of an ortholog database using the Semantic Web technology for integrative analysis of genomic data", PLoS ONE, 10, e0122802, 2015. (DOI:10.1371/journal.pone.0122802)	