DDBJ Sequence Read Archive

DDBJ Sequence Read Archive (DRA) は Roche 454 GS System®, Illumina Genome Analyzer®, Applied Biosystems SOLiD® System などの次世代シークエンサからの出力データのためのデータベースです。 DRA は International Nucleotide Sequence Database Collaboration (INSDC) のメンバーであり，NCBI Sequence Read Archive (SRA) と EBI Sequence Read Archive (ERA) との国際協力のもと，運営されています。 従来のキャピラリ式シークエンサからの出力データは DDBJ Trace Archive にご登録ください。

- 登録に必要なデータ
- 登録方法
- データの検索・ダウンロード
- DDBJ Read Annotation Pipeline でデータを解析

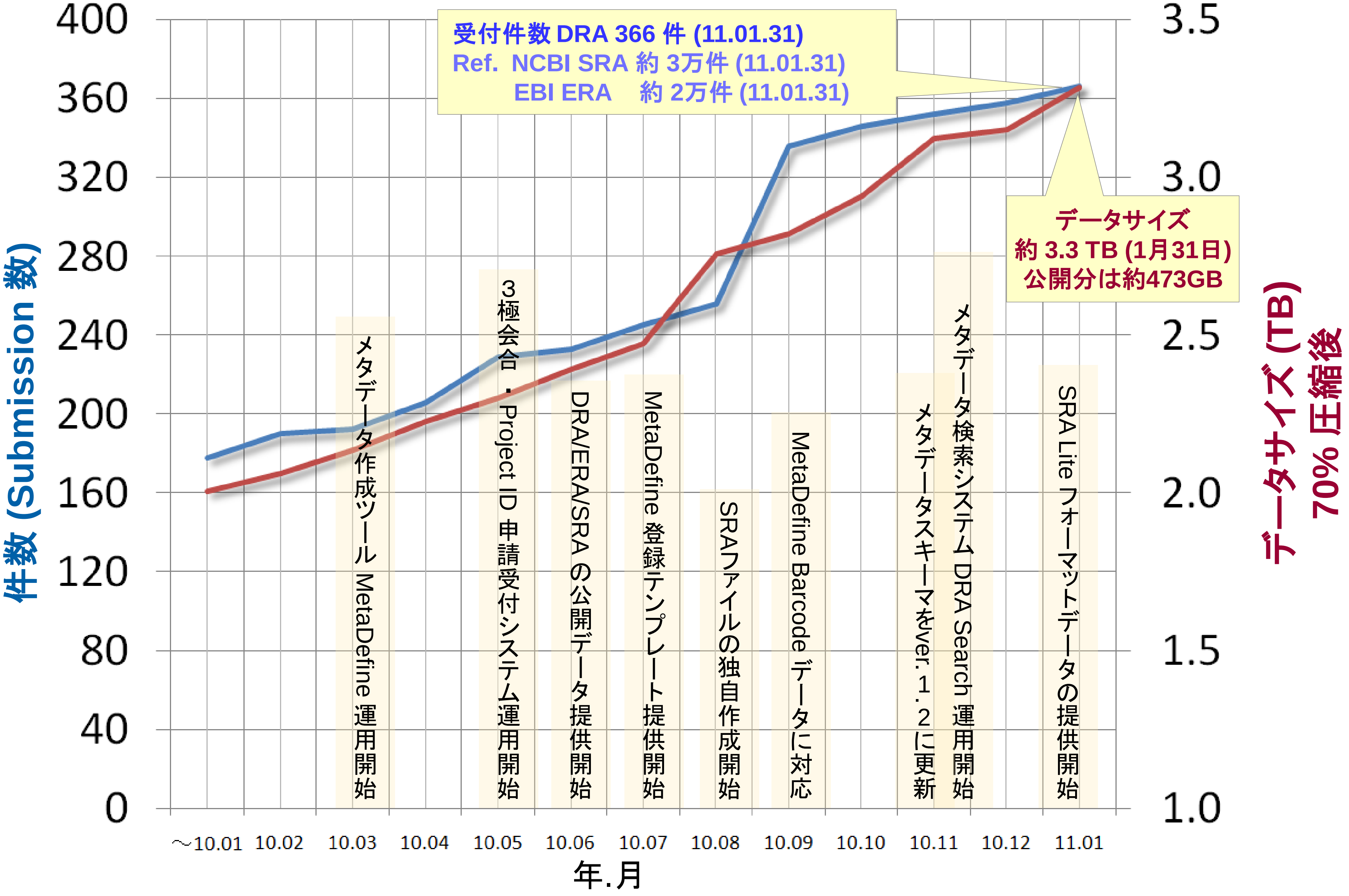
[2011-02-14] 2011-02-22 (火) 14:00-16:00 メンテナンス作業のため DRA 検索システム以外の登録システムと FTP データ提供サーバが利用できなくなります。何卒ご理解の程，よろしくお願いいたします。

DRA はライフサイエンス統合データベースプロジェクトの一部であり，科学技術振興機構のバイオインフォマティクス推進センターに支援されています。

Last modified: Feb. 14, 2011

DRA　DDBJ

DRA開発と受付の年間動向



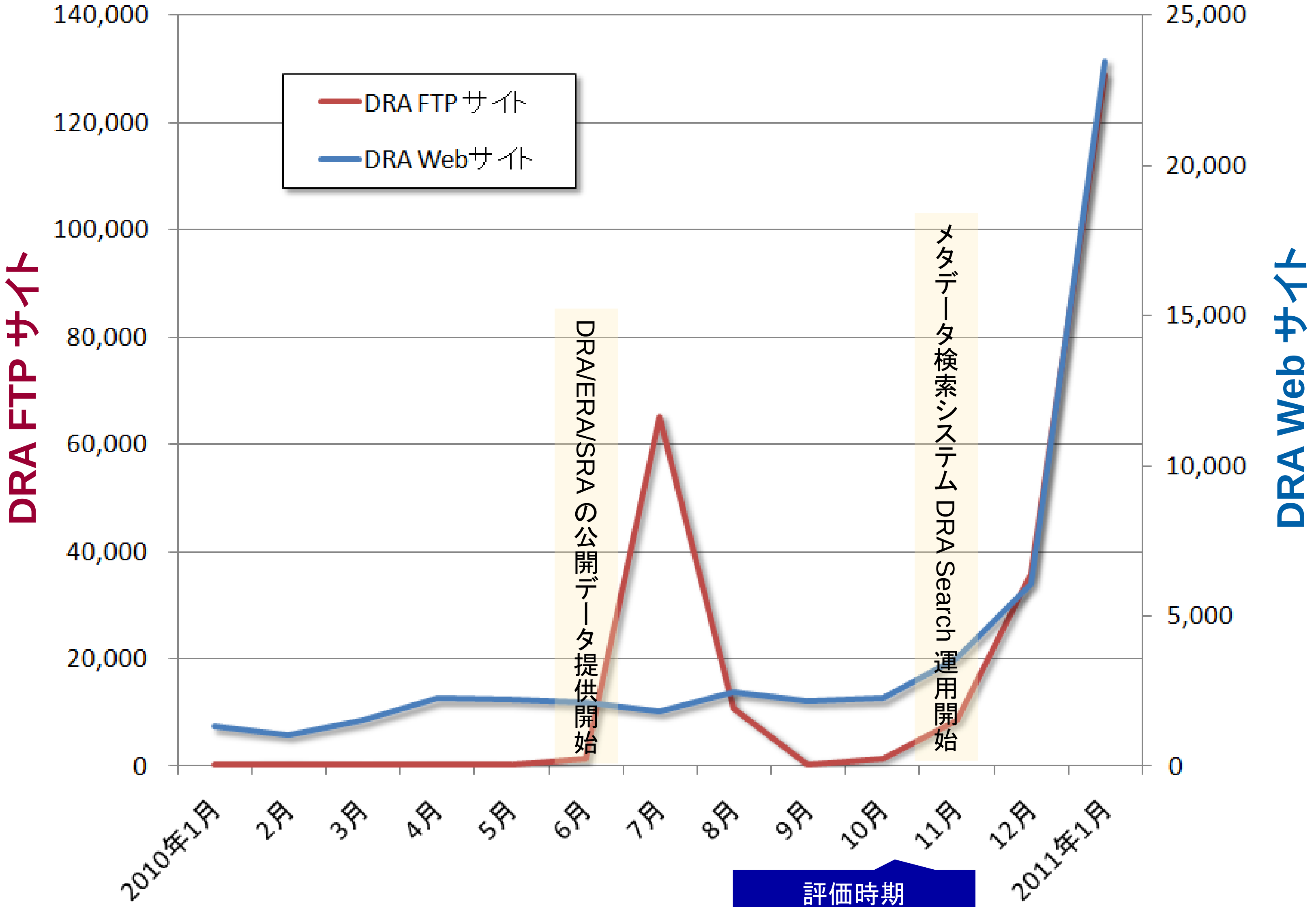
受付件数 DRA 366 件 (11.01.31)
Ref. NCBI SRA 約 3万件 (11.01.31)
EBI ERA 約 2万件 (11.01.31)

データサイズ
約 3.3 TB (1月31日)
公開分は約473GB

厳しい評価

- この程度のアクセス数のDBは存在意義が無いです。
- 重要なデータベースであるはずなのだが、アクセス数が少ないことは問題である。

DRAのアクセス動向（月変動）



評価時期
10月25日-11月14日

Announcements
2011-02-16: Sequence Read Archive (SRA) and Trace Archive repositories have been discontinued.

Due to budget constraints, NCBI will be discontinuing its Sequence Read Archive (SRA) and Trace Archive repositories for high-throughput sequence data. Closure of the databases will occur in phases. SRA and Trace will stop accepting some types of submissions in the coming weeks, and all submissions within the next 12 months. Over the next several months, NCBI will be working with staff from NIH Institutes that fund large-scale sequencing efforts to develop an approach for future access to and storage of the existing data. NCBI will continue to support and develop information resources for biological data derived from next-generation sequencing such as genotypes, common variations, rare variations, sequence assemblies and gene expression data. We therefore encourage the research community to continue submissions of these data to the applicable databases, including:

1. RNA-Seq and epigenomic data to GEO
2. Variants, genotypes, phased haplotypes, and polymorphisms to dbVar, dbGaP and dbSNP
3. Genomic assemblies to GenBank/WGS
4. Transcript assemblies to GenBank/TSA
5. 16S ribosomal RNA and other targeted locus survey assemblies to GenBank

NCBI expects new applications will continue to emerge for next generation technology. We are excited to work with the community to develop strategies for archiving other summary experimental measures that are informative, efficient, and valuable to the biomedical research community.

For further information about submissions, contact NCBI’s Help Desk.

http://trace.ncbi.nlm.nih.gov/Traces/sra/sra.cgi

EMBL-EBI will continue to support the Sequence Read Archive for raw data

Hinxton, 16 February 2011 - Because the rapid evolution of DNA sequencing technology has reduced the cost and time it takes to analyse genomes, an enormous amount of data is each day being submitted to, and retrieved from, the Sequence Read Archive (SRA), a public repository that preserves raw experimental DNA data.

Increasingly, journals and funding agencies are requiring that research leaders deposit their DNA data – for example from the 1000 Genomes Project – in the SRA, which is operated by the International Nucleotide Sequence Database Collaboration (INSDC). Until now, the SRA has been hosted and curated by INSDC partners the European Bioinformatics Institute (EMBL-EBI), the US National Center for Biotechnology Information (NCBI) and the DNA Databank of Japan (DDBJ).

This week, NCBI announced that owing to budget constraints they will need to phase out submissions to the SRA over the next 8 to 12 months. EMBL-EBI will continue to support the archive, and to collaborate with NCBI on many other projects. EMBL-EBI considers public, open-access nucleic acid sequence databases to be essential for research in the life sciences. As the fastest growing component of the European Nucleotide Archive (ENA), the SRA helps disseminate data from large- scale studies, relieving those who generate the data from the burden of storing and managing huge amounts of information. It also ensures that the wider scientific community can access and take advantage of the data for scientific discovery and application.

Work at EMBL-EBI on its SRA centres on the on-going development of scalable solutions to capture, preserve and disseminate next generation sequence data. As we face unprecedented data growth and ever-broadening applications for sequencing, this has become one of the most challenging areas of our work.
EMBL-EBI's approach to archiving sequence data depends on the deep involvement of the user community. We combine novel software compression techniques with judicious data reduction to achieve scalable storage of sequence data, and look to domain experts outside of our organisation to collaborate in developing the smoothest and most useful data submission and retrieval routes. The emerging ELIXIR framework for bioinformatics infrastructure in Europe provides an excellent forum to develop further this approach.

With community agreement on appropriate data reduction and the active community engagement to achieve smooth submission and retrieval pipelines, EMBL-EBI will continue to operate its SRA. But as all archives must justify their costs on the basis of future benefits, we will continue actively to consult and work with a broad community of scientists to understand the value of this and other life science archives. As a matter of principle and policy, any EMBL-EBI decision affecting the community in such a serious way would be communicated in good time to allow for orderly transitions.

EMBL-EBI will continue to work actively with NCBI on many projects, including the traditional assembled sequence archive (EMBL-Bank/GenBank/DDBJ)

http://www.ebi.ac.uk/ena/SRA_announcement_Feb_2011.pdf?utm_source=twitterfeed&utm_medium=twitter