

NCBI News, January 2010

Peter Cooper, PhD¹ and Dawn Lipshultz, MS²

Created: February 23, 2010; Updated: March 1, 2010.

New Databases and Tools

Bookshelf

New books added to the Bookshelf include *Advancing the Nation's Health Needs: NIH Research Training Programs*, *The Threat of Pandemic Influenza: Are We Ready?*, *Asbestos: Selected Cancers*, and *Implications of Nanotechnology for Environmental Health Research*. View these and other books on the Bookshelf Web service (www.ncbi.nlm.nih.gov/books/).

NCBI Publications

Six NCBI-related articles appear in the recent Database issue of *Nucleic Acids Research*.

1. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. **GenBank**. *Nucleic Acids Res.* 2010 Jan;38(Database issue):D46-51. Epub 2009 Nov 12. PMID: [19910366](https://pubmed.ncbi.nlm.nih.gov/19910366/)
2. Geer LY, Marchler-Bauer A, Geer RC, Han L, He J, He S, Liu C, Shi W, Bryant SH. **The NCBI BioSystems database**. *Nucleic Acids Res.* 2010 Jan;38(Database issue):D492-6. Epub 2009 Oct 23. PMID: [19854944](https://pubmed.ncbi.nlm.nih.gov/19854944/)
3. Ji L, Barrett T, Ayanbule O, Troup DB, Rudnev D, Muertter RN, Tomashevsky M, Soboleva A, Slotta DJ. **NCBI Peptidome: a new repository for mass spectrometry proteomics data**. *Nucleic Acids Res.* 2010 Jan;38(Database issue):D731-5. Epub 2009 Nov 26. PMID: [19942688](https://pubmed.ncbi.nlm.nih.gov/19942688/)
4. Sayers EW, Barrett T, Benson DA, Bolton E, Bryant SH, Canese K, Chetvernin V, Church DM, Dicuccio M, Federhen S, Feolo M, Geer LY, Helmberg W, Kapustin Y, Landsman D, Lipman DJ, Lu Z, Madden TL, Madej T, Maglott DR, Marchler-Bauer A, Miller V, Mizrachi I, Ostell J, Panchenko A, Pruitt KD, Schuler GD, Sequeira E, Sherry ST, Shumway M, Sirotkin K, Slotta D, Souvorov A, Starchenko G, Tatusova TA, Wagner L, Wang Y, John Wilbur W, Yaschenko E, Ye J. **Database Resources of the National Center for Biotechnology Information**. *Nucleic Acids Res.* 2010 Jan;38(Database issue):D5-16. Epub 2009 Nov 12. PMID: [19910364](https://pubmed.ncbi.nlm.nih.gov/19910364/)

¹ NCBI; Email: cooper@ncbi.nlm.nih.gov. ² NCBI; Email: lipshult@ncbi.nlm.nih.gov.

5. Shumway M, Cochrane G, Sugawara H. **Archiving next generation sequencing data**. Nucleic Acids Res. 2010 Jan;38(Database issue):D870-1. Epub 2009 Dec 3. PMID: 19965774
6. Wang Y, Bolton E, Dracheva S, Karapetyan K, Shoemaker BA, Suzek TO, Wang J, Xiao J, Zhang J, Bryant SH. **An overview of the PubChem BioAssay resource**. Nucleic Acids Res. 2010 Jan;38(Database issue):D255-66. Epub 2009 Nov 19. PMID: 19933261

Microbial Genomes

Twenty-two finished microbial genomes were added to the NCBI databases in January, 2010. The original sequence data files submitted to GenBank/EMBL/DDBJ are on the FTP site (<ftp.ncbi.nih.gov/genbank/genomes/Bacteria/>). The RefSeq provisional versions of these genomes are also available (<ftp.ncbi.nih.gov/genomes/Bacteria/>).

GenBank News

GenBank release 175.0 is available on the NCBI Web Service and FTP site. The current release incorporates sequence data as of December 15, 2009. A new GenBank release will be available in February. Release notes containing detailed information are available on the FTP site: <ftp.ncbi.nlm.nih.gov/genbank/gbrel.txt>.

Updates and Enhancements

RefSeq

RefSeq Release 39 is now available through the Entrez system and can be downloaded from the FTP site (<ftp.ncbi.nlm.nih.gov/refseq/release>). This full release incorporates genomic, transcript, and protein data available as of January 23, 2010. It includes 13,656,433 records from 10,171 different species and strains. Changes since the last release can be found in the release notes (<ftp.ncbi.nlm.nih.gov/refseq/release/release-notes/RefSeq-release39.txt>). More information on the RefSeq project is available on the RefSeq Homepage (www.ncbi.nlm.nih.gov/RefSeq/).

My NCBI

The My Bibliography portion of My NCBI now helps manage publication compliance with the NIH Public Access Policy. The 2010 January-February issue of the NLM Technical Bulletin includes an [article](#) that explains this new feature.

COBALT Improvements

The COBALT multiple sequence alignment service produces downloadable multiple-alignment output in the popular text formats FASTA plus gaps, Clustal, Phylip, Nexus, and ASN.1. This feature greatly increases the versatility of COBALT by allowing export of

results to alignment editors and other multiple sequence alignment tools. The [BLAST News](#) page has more information.

UniVec Database

A new version (build 5.2) of UniVec, NCBI's non-redundant database of vector sequences, is now available on the VecScreen Web service (www.ncbi.nlm.nih.gov/VecScreen/VecScreen.html).

The number of sequences in the database increased by three percent with this new build. The newly added sequences include 28 complete vector sequences, and 23 adaptor, primer and multiple-cloning site sequences.

The vector BLAST database, available from the BLAST area of the FTP site (ftp.ncbi.nih.gov/blast/db/), has also been updated to contain full-length versions of all sequences from GenBank that were used to build UniVec 5.2.

Exhibits

NCBI will have an exhibit booth at the American Association for the Advancement of Science annual meeting held February 18-22 in San Diego, CA.

Announce Lists and RSS Feeds

Eighteen topic-specific mailing lists are available which provide email announcements about changes and updates to NCBI resources including dbGaP, BLAST, GenBank, and Sequin. The various lists are described on the Announcement List summary page (www.ncbi.nlm.nih.gov/Sitemap/Summary/email_lists.html). To receive updates on the *NCBI News*, please see: www.ncbi.nlm.nih.gov/About/news/announce_submit.html

Twelve RSS feeds are now available from NCBI including news on PubMed, PubMed Central, NCBI Bookshelf, LinkOut, HomoloGene, UniGene, and NCBI Announce. Please see: www.ncbi.nlm.nih.gov/feed/

Send comments and questions about NCBI resources to: info@ncbi.nlm.nih.gov, or call 301-496-2475 between the hours of 8:30 a.m. and 5:30 p.m. EST, Monday through Friday.