

# NCBI News, June 2010

Peter Cooper, Ph.D.<sup>1</sup> and Dawn Lipshultz, M.S.<sup>2</sup>

Created: June 23, 2010.

## New Databases and Tools

### Selected Structures

A new Selected Structures filter is available with Entrez Structure (MMDB) search results. This feature provides a way of sorting results into subsets based on characteristics of the retrieved records. Available subsets are keyed to protein domain families, source organisms, the presence of specific molecular complexes, and the presence of links to literature from PubMed or PubMed Central. The Selected Structures filter links appear in a box in the upper right corner of the Structure search result pages. Clicking on any of the linked numbers produces a subset of the results with the listed property. A portion of the results of a search for “p53 tumor suppressor” in the Structure database with the Selected Structures feature is shown below. Clicking on the filter for those records with protein-DNA complexes (green arrow) narrows the results to the four structures that have bound DNA (lower panel).

---

<sup>1</sup> NCBI; Email: cooper@ncbi.nlm.nih.gov. <sup>2</sup> NCBI; Email: lipshult@ncbi.nlm.nih.gov.

The screenshot displays the NCBI Structure database interface. The main content area shows search results for 'P53 Core Dimer Bound To Dna' and related structures. The results are organized into a list with checkboxes, each containing a 3D structure visualization and associated metadata.

**Items 1 - 58 of 58** One page.

1: 2GEQ Related Structures, Literature, Domains, Chemicals, Other Links

Crystal Structure Of A P53 Core Dimer Bound To Dna [TranscriptionDNA]  
 Taxonomy: *Mus musculus*, synthetic construct  
 Proteins: 2; Nucleic acids: 2 (DNA); Chemicals: 2  
 modified: 2009/09/04; MMDB ID: 76523

2: 3GLU Related Structures, Literature, Domains, Chemicals, Other Links

Crystal Structure Of Human Sirt3 With Acecs2 Peptide [HydrolaseHYDROLASE  
 REGULATOR, EC: 3.5.1.-]  
 Taxonomy: *Homo sapiens*  
 Proteins: 2; Chemicals: 2  
 modified: 2009/09/17; MMDB ID: 73353

**Items 1 - 4 of 4** One page.

1: 2GEQ Related Structures, Literature, Domains, Chemicals, Other Links

Crystal Structure Of A P53 Core Dimer Bound To Dna [TranscriptionDNA]  
 Taxonomy: *Mus musculus*, synthetic construct  
 Proteins: 2; Nucleic acids: 2 (DNA); Chemicals: 2  
 modified: 2009/09/04; MMDB ID: 76523

2: 1BF5 Related Structures, Literature, Domains, Other Links

Tyrosine Phosphorylated Stat-1DNA COMPLEX [Gene RegulationDNA]  
 Taxonomy: *Homo sapiens*, synthetic construct  
 Proteins: 1; Nucleic acids: 2 (DNA)  
 modified: 2009/07/16; MMDB ID: 71645

3: 1TUP Related Structures, Literature, Domains, Chemicals, Other Links

Tumor Suppressor P53 Complexed With Dna [Antitumor ProteinDNA]  
 Taxonomy: *Homo sapiens*, synthetic construct  
 Proteins: 3; Nucleic acids: 2 (DNA); Chemicals: 1  
 modified: 2007/10/13; MMDB ID: 51571

4: 1TSR Related Structures, Literature, Domains, Chemicals, Other Links

P53 Core Domain In Complex With Dna [Antitumor ProteinDNA]  
 Taxonomy: *Homo sapiens*, synthetic construct  
 Proteins: 3; Nucleic acids: 2 (DNA); Chemicals: 1  
 modified: 2007/10/13; MMDB ID: 51561

**Selected Structures** Structure Count

Protein Domain Families

Families 42

P53: P53 DNA-binding domain 20

SIRT1 16

RPA1N 3

HECTc 2

ANK 1

... All 8 Families

Superfamilies 57

P53: P53 DNA-binding domain 20

Complexes

Protein-Protein 47

Protein-DNA 4

Protein-Chemical 39

Literature

PubMed 57

PMC 23

Taxonomy 58

*Homo sapiens*, (human), species, primates 37

Synthetic construct, species, other sequences 14

*Saccharomyces cerevisiae*, (baker's yeast), species, ascomycetes 11

*Danio rerio*, (zebrafish), species, bony fishes 3

*Mus musculus*, (house mouse), species, rodents 2

... All 10 Organisms

See the [structure help documentation](#) to read more about Selected Structures.

## New Nucleotide and Protein Pages

The [Nucleotide](#) and [Protein](#) pages have been updated with a new format and display settings. One important new feature is the ability to download coding sequence (CDS) regions as FASTA formatted sequences directly from the parent records. CDS downloads can be accessed using the new “Send” menu in the upper right of single or multiple record displays. See the [note](#) on the NCBI Facebook page for additional details.

## Taxonomy Links to Wikipedia

Over 52,000 NCBI Taxonomy pages now contain links to Wikipedia articles through the External Information Resources section (LinkOut). To see an example, visit the taxonomy record for *Saimiri oerstedii*, the Central American Squirrel monkey. You can follow the link to the [Wikipedia article](#) about this species. These new LinkOuts are provided through the [iPhylo](#) project.

## Microbial Genomes

Twenty-one finished microbial genomes were released in June 2010. The original sequence data files submitted to GenBank/EMBL/DDBJ are available 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 178.0 is available through the web and FTP sites. The current release includes information available as of June 11, 2010. The Release Notes provide more details: <ftp.ncbi.nih.gov/genbank/gbrel.txt>

## Updates and Enhancements

### Genome Workbench

Genome Workbench version 2.1.0 is available with numerous fixes and enhancements, giving researchers an even more powerful tool. Enhancements include an improved tree viewer used for phylogenetic tree analysis, integration of the MUSCLE alignment tool, and numerous updates to graphical views. [Release notes](#) for this version provide details on new features and bug fixes.

Genome Workbench is available for Windows, Linux, and Mac OS X and can be downloaded from the homepage: [www.ncbi.nlm.nih.gov/projects/gbench/](http://www.ncbi.nlm.nih.gov/projects/gbench/).

### PubMed Central Turns Ten

PubMed Central celebrated its 10<sup>th</sup> Anniversary this year. Created as a free, online archive in 2000, it has grown to contain two million articles from 37,000 journals, over 600 of which are full deposit journals. A complete history has been compiled in the May-June issue of the [NLM Technical Bulletin](#). NCBI celebrated the anniversary with a short symposium with speakers both from the NCBI and the outside community in May 2010.

### PubMed

Recent improvements to PubMed include an updated Advanced Search page that provides a better-integrated and more intuitive mechanism for constructing complex searches and the ability to select what data are sent through the E-mail option on the Send menu. The improved Advanced Search page and changes to the E-mail option are described in more detail in articles [e9](#) and [e10](#) in the May-June 2010 issue of the [NLM Technical Bulletin](#).

### MyNCBI

My NCBI now allows sign-in by partner organization accounts. The improvement is described in more detail in the recent [NLM Technical Bulletin](#).

## YouTube

Four How-to Tutorial videos have been added to NCBI's YouTube channel: [www.youtube.com/ncbinlm](http://www.youtube.com/ncbinlm). The following are the currently available instructional videos:

- Download a custom set of records.
- Obtain genomic sequence for and near a gene ([short](#) and [extended](#) versions).
- Retrieve all sequences for an organism.

## 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](http://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](http://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/](http://www.ncbi.nlm.nih.gov/feed/).

Users can also stay updated on NCBI's resources on Facebook ([www.facebook.com/ncbi.nlm](http://www.facebook.com/ncbi.nlm)) and Twitter ([twitter.com/NCBI](https://twitter.com/NCBI)).

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