

NCBI News, September 2015

First offering of NCBI NOW (Next generation sequencing Online Workshop) to begin October 13, 2015

Wednesday, September 30, 2015

From October 13th to October 23rd, NCBI will present the first iteration of NCBI NOW, a free online experience aimed at those new to next generation sequencing (NGS) analysis. Enrollment in this course is limited to the first 1,000 participants who sign up through the [ORAU Portal](#). Since enrollment is on a first-come, first-served basis, please only sign up for this educational opportunity if you will be able to participate fully.

NCBI NIH

NOW!

An online workshop that introduces next generation sequencing with a hands-on component in the cloud!

7 online lectures

self-guided

hands-on

DNaseq, BLAST, RNASeq, GEO, SRA

Learners will watch 6-7 videos (average video duration: 45-60 minutes) online during the first 7 days of the course. These videos will cover the basics of NGS data, preprocessing, quality control and alignment strategies for both DNA-Seq and RNA-Seq, as well as a brief discussion of downstream analysis. Additionally, we will demonstrate how to leverage BLAST tools for NGS analysis.

Next, participants will apply a selection of RNA-Seq alignment algorithms over three days (1-2 hours per day), mapping RNA-Seq data to [GRCh38](#) chromosome 20. Finally, participants will compare the results of these mappers for specific genes. Throughout the

course, participants will be able to post questions at [Biostars](#); experts from NCBI and elsewhere will be available online to answer questions.

Learners will emerge from the course equipped to map their own RNA-Seq or DNA-Seq data to the human genome, understand the options for downstream analysis, and use their understanding of the basic steps of data processing to interact more effectively with bioinformatician collaborators.

A note about registration: When registering for an account, you will be prompted for a partition. Please ensure that NCBI is selected, as this will provide access to the proper materials for this workshop. However, if you have already created an account without selecting this option, there is no need to create a second account. Your account has been reassigned to the NCBI partition.

September 30th NCBI Minute: Preview of NCBI at American Society of Human Genetics 2015

Thursday, September 17, 2015

On September 30th, NCBI staff will provide a quick overview of NCBI activities at this year's [American Society of Human Genetics](#) (ASHG) meeting, including previews of NCBI posters and presentations on tools and resources for clinical genetics, genomics and the assembly and annotation of the human genome.

NCBI will participate in the ASHG annual meeting in Baltimore, MD, October 6-10, 2015. Staff members will be at exhibit booth #2405, where attendees can get answers and provide input for the future development of NCBI human genome resources.

Date and time: Wednesday, September 30, 2015 12:00PM EDT

Registration URL: <https://attendee.gotowebinar.com/register/8493856336405913090>

After the live presentation, the webinar will be uploaded to the [NCBI YouTube channel](#). The webinar and any materials will also be archived on the [Webinars and Courses page](#), where you can also find information about future webinars.

"Create a Biosketch with SciENCv" webinar recording on YouTube

Wednesday, September 16, 2015

The recording of the July 30th webinar ([Announcement](#)) on SciENCv and Biosketch is available on [YouTube](#). The webinar shows you how to use SciENCv to maintain your scientific record and generate the new NIH Biosketch; this [NCBI Insights blog post](#) complements the webinar.

To receive alerts about new videos ranging from quick tips to full webinar presentations, subscribe to the [NCBI YouTube channel](#).

October 2nd webinar - LinkOut: Linking to datasets, databases and more

Wednesday, September 16, 2015

On October 2nd, NCBI staff will present a webinar on [LinkOut](#), an NCBI service that allows you to link directly from NCBI databases to a wide range of relevant information beyond the NCBI systems. This webinar will provide an overview of the service and highlight resources that participate in LinkOut, with a special emphasis on resources beyond full text articles, including databases, datasets and research tools.

If you use NCBI databases, produce databases, datasets or resources, or are a librarian supporting research and data science, this webinar is for you.

Date and time: Friday, October 2, 2015 12:00PM EDT

Registration URL: <https://attendee.gotowebinar.com/register/5533821500870613249>

After the live presentation, the webinar will be uploaded to the [NCBI YouTube channel](#). The webinar and any materials will also be archived on the [Webinars and Courses page](#), where you can also find information about future webinars.

New NCBI Insights blog post: "Finding Chemical Probes & Modulators - The Hunt for New Chemical Reagents and Medicines"

Friday, September 11, 2015

The [latest blog post on NCBI Insights](#) will show you how to use PubChem to find chemicals that bind to a particular gene or protein target and download a table with that information. In [last week's post](#), we showed you how to do the reverse - find and download a table of gene and protein targets for a particular chemical.

[NCBI Insights](#) is the official NCBI blog, where we share science features, quick tips and updates on what's new at NCBI.

NCBI to hold fourth offering of "A Librarian's Guide to NCBI"

Thursday, September 10, 2015

The NCBI, in partnership with the [National Library of Medicine Training Center \(NTC\)](#), will once again offer the [Librarian's Guide to NCBI](#) course on the NIH campus, March 7 - 11, 2016 ([Announcement](#)). If you are a medical or science librarian in the United States who offers bioinformatics education and support services, or are planning to offer such services in the future, please join us for this intensive exploration of modern molecular biology, genetic data and other biomedical data as represented at the NCBI.

The course explains how and why these data are generated, their importance in modern biomedical research, and how to access them through the NCBI website. The [online application](#) is now open. The application deadline is **September 14, 2015**. This is a combined application for the prerequisite online *Fundamentals of Bioinformatics and Searching* and the five-day in-person course. If you have already completed the *Fundamentals* class, please let us know on the application. Anyone who has previously taken *Fundamentals* is eligible to apply for the in-person course in March at NIH.

Prerequisite: *Fundamentals of Bioinformatics and Searching*

Because of the fast-paced and intense nature of the course, all applicants for *A Librarian's Guide* must successfully complete the online *Fundamentals of Bioinformatics and Searching* class offered through the NTC. The *Fundamentals* class is an introduction to molecular biology and bioinformatics taught by Diane Rein, Ph.D, MLS that provides essential background for the in-person course.

For those who have not yet taken the online course, a special section of *Fundamentals* will be offered in the fall of 2015 from **October 26 - December 11**, with classes not held the week of Thanksgiving.

September 16th NCBI Minute: Accessing the Human Genomics Standard Data (Genome in a Bottle) at NCBI

Wednesday, September 09, 2015

In the next NCBI Minute, we will show you how to access the Genome in a Bottle dataset from the NCBI site and present potential use cases.

Date and time: Wednesday, September 16, 2015 1PM EDT

Registration URL: <https://attendee.gotowebinar.com/register/1790003099972245250>

The [Genome in a Bottle Consortium](#) aims to provide well-characterized reference materials for human clinical sequencing and spur development and optimization of genomics technology and bioinformatics.

After the live presentation, the webinar will be uploaded to the [NCBI YouTube channel](#). The webinar and any materials will also be archived on the [Webinars and Courses page](#), where you can also find information about future webinars.

New NCBI Insights blog post: "Identifying Chemical Targets - Finding Potential Cross-Reactions and Predicting Side Effects"

Friday, September 04, 2015

The latest NCBI Insights [blog post](#) will show you how to use PubChem to download a table of gene and protein targets for a particular chemical, which can help to find potential

cross reactions and side effects. Next week, we will show you how to do the opposite - download a table of chemicals that bind to a particular gene or protein target.

[NCBI Insights](#) is the official NCBI blog, where we share science feature stories, quick tips and updates on what's new at NCBI, including [PubMed Labs](#).

RefSeq Release 72 is now available

Thursday, September 03, 2015

RefSeq Release 72 is now available online, on the [FTP site](#), and through NCBI's programming utilities. The full release incorporates genomic, transcript, and protein data available, as of August 27, 2015 and includes 79,189,847 records, 51,933,925 proteins, 12,321,036 RNAs and sequences from 54,937 organisms. More information can be found in the [release notes](#).

Release 72 includes over 43,000 Archaeal and Bacterial RefSeq genomes that were annotated since July 29, 2015 using one version of the prokaryotic genome annotation pipeline software (version 3.0). This re-annotation increases overall consistency across the dataset because all genomes were annotated in a short window of time using the same software and protein alignment data.

This release also incorporates modified policies related to protein annotation that address concerns with very short partial proteins (fragments), partial proteins where the neighboring sequence is good (a/c/g/t bases and no gaps), and proteins that incorporated many 'X' amino acid residues due to low quality genome sequence for the CDS region. These changes in protein annotation resulted in the suppression of 6.4 million non-redundant (WP_) protein records., the majority of which were partial proteins (>4 million).

At the same time, we created 3.8 million new protein records as we added approximately 12,000 new prokaryotic genomes to the RefSeq collection. An average of 316 new protein records were added for each of the 12,000 new genomes.; the remaining protein complement per genome is based on pre-existing non-redundant protein records. For more information about the RefSeq project, pipelines and data, please take a look at the [RefSeq homepage](#).

HIV-1 interaction datasets in Gene updated

Tuesday, September 01, 2015

NCBI has added data provided by the Southern Research Institute (SRI) to the HIV-1 interaction datasets available in [Gene](#).

The [protein interactions dataset](#) now has:

- 7,567 interactions;
- 15,074 interaction descriptions;

- 3,623 proteins encoded by 3,582 human genes;
- and 6,610 publications.

The [replication interactions dataset](#) now has:

- 1,298 interactions;
- 1,369 interaction descriptions;
- 1,298 proteins encoded by 1,298 human genes;
- and 94 publications.

Data are also available at the [RefSeq HIV-1 website](#) and the [GeneRIF FTP site](#).

Genome Workbench 2.9.5 now available

Tuesday, September 01, 2015

[Genome Workbench 2.9.5](#) is available, as of August 27th. New features include added support for Mac OS 10.10 (Yosemite) and an added ruler for protein coding regions in graphical sequence view.

For the full list of fixes, improvements and features, see the [Genome Workbench release notes](#).

dbSNP build 145 (pig, chicken, sorghum, gibbon) now available

Tuesday, September 01, 2015

dbSNP build 145 is now available through the integrated NCBI Entrez system and [FTP](#). This release includes data for pig, chicken, sorghum and gibbon. Build 145 provides more than 195 million submitted variants and 84 million reference variants for 4 species. To see complete build statistics, visit the [SNP summary page](#).