NCBI News, June 2011

Peter Cooper, Ph.D.¹ and Rana Morris, Ph.D.² Created: May 13, 2011; Updated: June 30, 2011.

Featured Resource: Re-designed PopSet

NCBI's PopSet database of related sequences and alignments from phylogenetic, population, mutation, and ecosystem studies has been completely redesigned and now features an embedded graphical alignment and better integration of related data from other PopSets and other Entrez databases. The new pages also include on-the-fly analysis with BLAST and Tree View.

The PopSet Record View

The PopSet record view is now fully integrated with the updated Entrez system and can be addressed simply with the PopSet database name and the identifier as shown below.

http://www.ncbi.nlm.nih.gov/popset/298351991

The record display shown in Figure 1 consists of up to three sections: the study details showing the article reporting the current set; a list of the sequence records in the set; and, when available, the submitted alignment displayed in the embedded Graphical Sequence Viewer (GSV), now also appearing in Entrez Gene and SNP record views. The PopSet embedded alignment view shows the alignment portion of the full GSV display of the master or top sequence in the multiple-alignment. Clicking on the "Open full-view" link opens the GSV nucleotide view of the top sequence showing the detailed alignment tracks.

As in the other Entrez databases, the "Display Setting" menu controls the format of the records displayed; the "Send to" menu manages saving data, shown in Figure 2. Display options are similar to those available for the Nucleotide database and include the standard sequence formats such as FASTA and GenBank. The sequence record formats are presented within the PopSet display rather than by linking to the sequence database.

The "Send to" menu can send data to the Entrez clipboard, Collections in a My NCBI account, or to a file on the local computer. The file saving format options include the standard sequence formats, popular multiple alignment formats – FASTA plus gap,

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

NLM Citation: Cooper P, Morris R. NCBI News, June 2011. 2011 May 13 [Updated 2011 Jun 30]. In: NCBI News [Internet]. Bethesda (MD): National Center for Biotechnology Information (US); 1991-2012.

	s: 🕑 PopSet							Send	<u>to:</u> 🖂						_		
										Analy	ze this d	ata set					
Carnivo	ra apolipopr	otein P) dene	narf	ial co	le			Run BL	AST align	ment					
		oteni D		y gene,	Pari					Tree Vi	ew						
PopSet: 2983																	
GenBank FAS	STA									A							
												ng this d					
<u>So to:</u> 🖂												of diversif Carnivora					
Study Det	ails											ene [Mol Ph			010]		
•	d timing of dive	relficatio	n of the m	ammalian	ord		uvora i	nforres									
	ple nuclear gen			ammanan	ord		nvora I	merrec									
	rphy,W.J., Koepfli,K.			oo.J.W., Wav	/ne.R I	C and O	Brien.S.	J.				s from th					
	logenet. Evol. 56:(1)49		.,, Diug	, may			2			Lauras	atheria RA	SA2 gene,	partial :	sequence	э.		
MID: 2013822											ra recomb	ination acti	vation p	orotein 2			
Go to: 🖂												tyrosine ph	oenhate		tor		
		Allements								Gamiyo	na protein	groane pr	oopliate	1001000	101		
Seguence	s in this data	Alignme	m														
equence	1																
U930905.1	Ursus americanus ar	253 - 281 (2	29 bases shown)											Op	en Full Vie	w • ?
<u>5U930905.1</u> 5U930904.1	Ursus americanus ap Ailuropoda melanolei	253 - 281 (2 ම, ල, m	29 bases shown)											<u>Op</u>	en Full Vie	× •?
GU930905.1 GU930904.1 GU930903.1	Ursus americanus ar Ailuropoda melanole Ailurus fulgens apolir		29 bases shown) Seq. S First	A	gnment									Op Last		
3U930905.1 3U930904.1 3U930903.1 3U930902.1	Ursus americanus ap Ailuropoda melanole Ailurus fulgens apolip Odobenus rosmarus	ତ୍ତ୍ଲ	29 bases shown			gnment 253		260			270			26	Last		2
6U930905.1 6U930904.1 6U930903.1 6U930902.1 6U930901.1	Ursus americanus ap Ailuropoda melanole Ailurus fulgens apolip Odobenus rosmarus Mirounga angustirost	Q Q m Description	29 bases shown 🔶 → Mar		253 A	53 A T C	ССТ		 Т А Т	G T A		A T T	, , T C I	CAG	Last 11 6 28	Seq. End	20 Seq. L 933
SU930905.1 SU930904.1 SU930903.1 SU930902.1 SU930901.1 SU930900.1	Ursus americanus ar Ailuropoda melanole Ailurus fulgens apolip Odobenus rosmarus Mirounga angustiros Mydaus marchei apo	⊙ ⊙ no Description	29 bases shown 🔶 → Mar		253 A 253 A	53 A T C	ССТ		тат тат тат	G T A G T A G T A G T A	A A A G A A A G	A T T A T T A T T A T T	тсі тсі тсі		Last 1 6 28 6 28	Seq. End 1 933 1 933 1 933	2 Seq. L 933 933 933
SU930905.1 SU930904.1 SU930903.1 SU930902.1 SU930901.1 SU930900.1 SU930909.1	Ursus americanus ap Ailuropoda melanole Ailurus fulgens apolip Odobenus rosmarus Mirounga angustirost	Q Q @ Description GU9308864.1 GU930886.1 GU930886.1 GU930886.1	29 bases shown 🔶 → Mar	Seq. S First	253 A 253 A 253 A 253 A	A T C A T C A T C A T C	ССТ			G T A G T A G T A				CAG CAG CAG	Last G 28 G 28 G 28 G 28 G 28	Seq. End 1 933 1 933 1 933 1 933	20 Seq. L 933 933 933 933
EU930905.1 EU930904.1 EU930903.1 EU930902.1 EU930901.1 EU930900.1 EU930899.1 EU930898.1	Ursus americanus ar Ailuropoda melanole Ailurus fulgens apolit Odobenus rosmarus Mirounga angustiros Mydaus marchei apo Conepatus leuconotu	Q Q m Description GU930884.1 GU930885.1 GU930886.1 GU930886.1 GU930888.1 GU930888.1 GU930888.1	29 bases shown	Seq. S First	253 A 253 A 253 A 253 A 253 A 252 A 252 A	A T C A T C A T C A T C A T C A T C	ССТ	260 T	T A T T A T	G T A G T A G T A G T G G T G	A A A G A A A G		тс	CAG CAG CAG CAG CAG CAG	Last 1 6 28 8 6 28 8 6 28 8 6 28 8 6 28 8 6 28 8 6 28 8 6 28 8 6 28 8 8 8 8 8 8 8 8 8 8 8 8 8	Seq. End 1 933 1 933 1 933 1 933 0 932 0 932	2 Seq. L 933 933 933 933 933 932 932
5U930905.1 5U930904.1 5U930903.1 5U930902.1 5U930901.1 5U930900.1 5U930899.1 5U930898.1 5U930897.1	Ursus americanus ar Ailuropoda melanoler Ailurus fulgens apolir Odobenus rosmarus Mirounga angustirost Mydaus marchei apo Conepatus leuconotu Spilogale putorius ap	Q. Q. m Description GU930884.1 GU930886.1 GU930886.1 GU930887.1 GU930889.1 GU930889.1 GU930889.1	29 bases shown	Seq. S First	253 A 253 A 253 A 253 A 252 A 252 A 252 A 252 A	853 A T C A T C A T C A T C A T T A T T		260 T T T T T T	, , , , Т А Т Т А Т	G T A G T A G T A G T G G T G G T G	A A A G A A A G	ATT		C A G C A G C A G C A G C A G C A G C A G	Last G 28 G 28	Seq. End 1 933 1 933 1 933 1 933 0 933 0 933 0 933 0 933	8eq. L 933 933 933 933 933 932 932 932 932 932
SU930905.1 SU930904.1 SU930903.1 SU930903.1 SU930902.1 SU930900.1 SU930898.1 SU930898.1 SU930897.1 SU930895.1	Ursus americanus ar Ailuropoda melanole Ailurus fulgens apoliț Odobenus rosmarus Mirounga angustirost Mydaus marchei apo Conepatus leuconot. Spilogale putorius ap Mephitis mephitis ap	Q. Q. m Description GU930886.1 GU930886.1 GU930886.1 GU930887.1 GU930887.1 GU930889.1 GU930889.1 GU930889.1 GU930889.1 GU930889.1 GU930889.1 GU930889.1 GU930889.1 GU930889.1	29 bases shown 4	Seq. S First	253 A 253 A 253 A 253 A 252 A 252 A 252 A 252 A 252 A 252 A 252 A 253 A	A T C A T C A T C A T C A T C A T C A T T A T T A T T A T T A T T		260 T T T T T T	T A T T A T	G T A G T A G T A G T G G T G G T G		A T T A T T A T T	T C 0 T C 0 T C 0 T C 0	C A G C A G	Last G 28 G 28	Seq. End 1 933 1 933 1 933 1 933 0 933 0 933 0 933 0 933 1 933	2 Seq. L 933 933 933 933 932 932 932 932 932 932
SU930905.1 SU930903.1 SU930903.1 SU930902.1 SU930900.1 SU930900.1 SU930898.1 SU930898.1 SU930896.1 SU930895.1 SU930895.1	Ursus americanus ar Ailuropoda melanole Ailurus fuigens apolig Odobenus rosmarus Mirounga angustirost Mydaus marchei apo Conepatus leuconotu Spilogale putorius ap Mephitis mephitis ap Urocyon cinereoarge Nyctereutes procyon Genetta genetta apo	Q. Q. m GU930884.1 GU930885.1 GU930886.1 GU930886.1 GU930886.1 GU930889.1 GU930892.1 GU930892.1 GU930892.1 GU930893.1	29 bases shown 29 bases shown 40 bit 50 bit	Seq. S First	253 A 253 A 253 A 253 A 252 A 252 A 252 A 252 A 252 A 253 A 253 A	A T C A T C A T C A T C A T C A T T A T T A T T A T T A T T A T T		260 T T T T T T	, , , , T A T T A T	G T A G T A G T A G T G G T G		A T T A T T A T T A T T A T T A T T A T T		C A G C A G	Last 11 4 6 6 28 8 8 8 8 8 8 8 8 8 8 8 8 8	Seq. End 1 933 1 933 1 933 0 932 0 933 0 933 1 935 1 935	2 Seq. L 933 933 933 932 932 932 932 932
SU930905.1 SU930904.1 SU930903.1 SU930902.1 SU930900.1 SU930990.1 SU930898.1 SU930898.1 SU930895.1 SU930895.1 SU930895.1 SU930894.1 SU930893.1	Ursus americanus ap Aliuropoda melanole Aliurus fulgens apolit Odobenus rosmarus Mirounga angustrose Mydaus marchei apo Conepatus leuconotu Spilogale putorius ap Mephitis mephitis ap Urocyon cinereoarge Nyctereutes procyon Genetta genetta apo Civetitcis civetta apo	Q Q m Description GU9308684.1 GU930865.1 GU930866.1 GU930886.1 GU930887.1 GU930886.1 GU930892.1 GU930892.1 GU930893.1 GU930893.3 GU930892.3 GU930893.3 GU93084.4 GU930	29 bases shown 29 bases shown 40 bit 50 bit	Seq. S First	253 A 253 A 253 A 252 A 252 A 252 A 252 A 252 A 253 A 253 A 253 A	A T C A T C A T C A T C A T C A T T A T T		260 T T T T T T T T	 T A T T A T	G T A G T A G T A G T G G T G G T G G T G G T G G T G G T G A T G		A T T A T T		C A G C A G	Last Last G 286 G 286 G 286 G 286 G 286 G 286 G 286 G 286 G 286 G 286 G 286 G 286 G 286 G 286 G 286 G 286 G 28	Seq. End 1 933 1 933 1 933 0 933 0 933 0 933 1 935 1 93 1 935 1 935	2 5eq. L 933 933 933 933 932 932 932 932
U930905.1 SU930904.1 SU930903.1 SU930902.1 SU930901.1 SU930900.1 SU930898.1 SU930895.1 SU930895.1 SU930895.1 SU930895.1 SU930895.1 SU930895.1 SU930895.1	Ursus americanus ar Ailuropoda melanole Ailurus fulgens apolit Odobenus rosmarus Mirounga angustirost Mydaus marchei apo Conepatus leuconotu Spilogale putorius ap Mephilis mephilis ap Urocyon cinereoarge Nyctereutes procyon Genetta genetta apod Civetticits civetta apod Fossa fossana apolit	Q Q m Description GU930864.1 GU930865.1 GU9308867.1 GU9308867.1 GU930886.1 GU930890.1 GU930893.1 GU930893.1 GU930894.1 GU930894.1 GU930894.1 GU930894.1 GU930894.1	29 bases shown 29 bases shown 40 bit 50 bit	Seq. S First	253 A 253 A 253 A 253 A 252 A 252 A 252 A 252 A 253 A 253 A 253 A 253 A 253 A 253 A 253 A	A T C A T T A T T		260 T T T T T T T T	 T A T T A T A T T A	G T A G T A G T G G T G G T G G T G G T G G T G G T G G T G G T G G T G G T G G T G G T G G T G		ATT ATT ATT ATT ATT ATT ATT ATT ATT	T C C T C C	C A G C A G	Last 1 6 266 6 266 6 266 6 286 6 286	Seq. End 1 933 1 933 1 933 0 932 0 933 0 933 0 933 1 93 1 933 1 933	20 303 303 303 933 933 932 932 933 933 933 93
U930905.1 SU930904.1 SU930902.1 SU930901.1 SU930901.1 SU930900.1 SU930899.1 SU930897.1 SU930897.1 SU930895.1 SU930895.1 SU930894.1 SU930893.1 SU930893.1 SU930893.1	Ursus americanus ap Ailuropoda melanole Ailurus fulgens apolig Odobenus rosmarus Mirounga angustros Mydaus marchei apo Conepatus leuconotu Spilogale putorius ap Urocyon cinereoarge Nyctereutes procyon Genetta genetta apo Civettictis civetta apo Fossa fossana apolig Rhynchogale melleri	Q Q m Description GU9308864 1 GU9308861 GU9308861 GU9308861 GU9308891 GU9308821 GU9308821 GU9308821 GU9308821 GU9308841 GU9308841 GU9308841	29 bases shown	Seq. S First	253 A 253 A 253 A 252 A 252 A 252 A 252 A 253 A 253 A 253 A 253 A 253 A 253 A 253 A 253 A 253 A	A T C A T T A T T		260 T T T T T T T T	TAT	G T A G T A G T A G T G G T G G T G G T G G T G G T G G T G A T G A T G G C C G C C			T C C T C C		Last 1 - - - - - - - - - - - - -	Seq. End 1 933 1 933 1 933 0 933 0 933 0 933 0 933 1 93 1 933 1 933	2033 933 933 932 932 932 932 932 933 933 9
2U930905.1 SU930904.1 SU930902.1 SU930902.1 SU930900.1 SU930900.1 SU930898.1 SU930898.1 SU930895.1 SU930895.1 SU930895.1 SU930893.1 SU930892.1 SU930892.1 SU930892.1	Ursus americanus ar Aliuropoda melanole Aliurus fulgens apolit Odobenus rosmarus Mirounga angustrose Mydaus marchei apo Conepatus leuconotu Spilogale putorius ap Mephitis mephitis ap Urocyon cinereoarge Nyctereutes procyon Genetta genetta apo Civetitcis civetta apo Fossa fossana apolit Rhynchogale melleri Ichneumia albicauda	Q. Q. Bescription GU9308641 GU9308651 GU9308861 GU3308861 GU330881 GU3308861 GU3308961 GU3308961 GU3308961 GU3308961 GU3308961 GU3308961 GU3308961 GU3308961 GU3308961	29 bases shown	Seq. S First	253 A 253 A 253 A 252 A 252 A 252 A 252 A 253 A	A T C A T T A T T C T		T T T T T T	T A T T A T T A T	G T A G T A G T A G T G G T G G T G G T G G T G G T G G T G G T G G T G G T G G T G G C C G C C G C C G C C G C C		A T T A T T			Last G 286 G 2	Seq. End 1 933 1 933 1 933 1 933 0 933 0 933 0 933 1 933	2000 200 2000 2
U30905.1 U30904.1 U30909.1 U30909.1 U30909.1 U30909.1 U30909.1 U30909.1 U30909.1 U30908.1 U3008.1 U30908.	Ursus americanus ar Ailuropoda melanoler Ailurus fulgens apolit Odobenus rosmarus Mirounga angustirost Mydaus marchei apo Conepatus leuconotu Spilogale putorius ap Mephitis mephitis ap Urocyon cinereoarge Nyctereutes procyon Genetta genetta apod Civettictis civetta apo Fossa fossana apolit Rhynchogale melleri Ichneumia albicauda	Q. Q. 00 Description 01/200864.1 01/200865.1 GUI930085.1 GUI930085.1 01/200861.1 GUI330085.1 GUI330085.1 01/20085.1 GUI330085.1 GUI330085.1 01/20085.1 GUI330085.1 GUI330085.1 GUI330085.1	19 bases shown \$\$	Seq. S First	253 A 253 A 253 A 252 A 252 A 252 A 252 A 252 A 253 A	A T C A T C A T C A T C A T C A T C A T C A T C A T C A T		T T T T T T	T A T T A T T A T T A T T A T	G T A G T A G T A G T G G T G G T G G T G G T G G T G G T G A T G G C C G C C G C C G C C G C C G C C G C C	а а а а а а а а а а а а а а а а а а а	A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T			Last 	Seq. End 1 933 1 933 1 933 0 933 0 933 0 933 0 933 1 933	2033 933 933 933 933 932 932 932 933 933 9
SU330905.1 SU330904.1 SU330904.1 SU330901.1 SU330901.1 SU330981.1 SU330881.1 SU330884.1 SU330884.1 SU330881.1 SU330881.1 SU330881.1 SU330884.1 SU33084.1 SU3308	Ursus americanus ap Aliuropoda melanole Aliurus fulgens apolit Odobenus rosmarus Mirounga angustrost Mydaus marchei apo Conepatus leuconotu Spilogale putorius ap Urocyon cinereoarge Nyctereutes procyon Genetta apod Civettictis civetta apo Fossa fossana apolit Rhynchogale melleri Ichneumia albicauda Helogale parvula apo Suricata suricatta ap	Q. Q. (Q. 100) Description CU330864.1 CU330865.1 CU330867.1 CU33087.1	19 bases shown 1	Seq. S First	253 A 253 A 253 A 252 A 252 A 252 A 252 A 253 A	-53 - A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T C T C T A T A T A T A T A T A T A T A T A T A T A T A T A T		T T T T T T	T A T T A T T A T T A T T A T T A T T A T	G T A G T A G T A G T G G T G G T G G T G G T G G T G G T G A T G G C C G C C G C C G C C G C C G C C G C A G C A	а а а а а а а а а а а а а а а а а а а	A T T A T T			Last 	Seq. End 1 933 1 933 1 933 0 933 0 933 0 933 0 933 1 933	2033 933 933 933 933 932 932 932 933 933 9
SU330905.1 3U330904.1 3U330904.1 3U330902.1 3U330902.1 3U330909.1 3U330898.1 3U330898.1 3U330898.1 3U330898.1 3U330898.1 3U330898.1 3U330888.1 3U330888.1	Ursus americanus ar Aliuropoda melanoler Aliurus fuigens apolit Odobenus rosmarus Mirounga angustiros Mydaus marchei apo Conepatus leuconotu Spilogale putorius ar Mephitis mephitis ap Urocyon cinereoarge Nyctereutes procyon Genetta genetta apo Civeticitis civetta apo Fossa fossana apolit Rhynchogale melleri Ichneumia albicauda Helogale parvula apo Suricata suricatta apo Suricata ano angolit	Q. Q. 00 Description 01/200864.1 01/200865.1 GUI930085.1 GUI930085.1 01/200861.1 GUI330085.1 GUI330085.1 01/20085.1 GUI330085.1 GUI330085.1 01/20085.1 GUI330085.1 GUI330085.1 GUI330085.1	19 bases shown 1	Seq. S First	253 A 253 A 253 A 252 A 252 A 252 A 252 A 253 A	A T C A T T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T T A T T A T T T A T A T T A T T A		T T T T T T	T A T T A T	G T A G T A G T A G T G G T G G T G G T G G T G G T G G T G G C C G C C G C C G C C G C C G C A G C A G C A	ан а	A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T			Last 	Seq. End 1 933 1 933 1 933 1 933 0 933 0 933 0 933 1 933	2033 933 933 933 933 932 932 932 932 932 9
SU330905.1 SU330904.1 SU330904.1 SU330902.1 SU330902.1 SU330901.1 SU330895.1 SU330895.1 SU330895.1 SU330895.1 SU330895.1 SU330895.1 SU330895.1 SU330895.1 SU330895.1 SU330895.1 SU330885.1 SU330885.1 SU330885.1	Ursus americanus ar Ailuropoda melanole Ailurus fulgens apolit Odobenus rosmarus Mirounga angustirost Mydaus marchei apo Conepatus leuconotu Spilogale putorius ap Mephitis mephitis ang Urocyon cinereoarge Nyctereutes procyon Genetta genetta apod Civetictis civetta apo Civetictis civetta apol Rhynchogale melleri Ichneumia albicauda Helogale parvula apo Suricata suricata apol Panthera onca apolip Leopardus pardalis a	Q Q Q M Description GU3308841 GU3308851 GU3308851 GU330887 GU330897 GU3308	19 bases shown 1	Seq. S First	253 A 253 A 253 A 252 A 252 A 252 A 252 A 253 A	A T C A T T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T T A T T A T T T A T A T T A T T A		T T T T T T	T A T T A T	G T A G T A G T A G T G G T G G T G G T G G T G G T G G T G G C C G C C G C C G C C G C C G C A G C A G C A	ан а	A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T			Last Last Last Last Last Last C 28 C	Seq. End 1 933 1 933 1 933 1 933 0 933 0 933 0 933 1 933	2033 933 933 933 933 932 932 932 932 932 9
SU930905.1 SU930904.1 SU930903.1 SU930902.1 SU930901.1 SU930901.1 SU930898.1 SU930898.1 SU930898.1 SU930897.1 SU930896.1	Ursus americanus ar Aliuropoda melanoler Aliurus fuigens apolit Odobenus rosmarus Mirounga angustiros Mydaus marchei apo Conepatus leuconotu Spilogale putorius ar Mephitis mephitis ap Urocyon cinereoarge Nyctereutes procyon Genetta genetta apo Civeticitis civetta apo Fossa fossana apolit Rhynchogale melleri Ichneumia albicauda Helogale parvula apo Suricata suricatta apo Suricata ano angolit	Q. Q. Description GUI300841 GUI300851	19 bases shown \$\vert \$\vert Mar 1 \$\vert 1 \$\vert	Seq S First	253 A 253 A 253 A 252 A 252 A 252 A 252 A 253 A	A T C A T T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T T A T T A T T T A T A T T A T T A		T T T T T T	T A T T A T	G T A G T A G T A G T G G T G G T G G T G G T G G T G G T G G C C G C C G C C G C C G C C G C A G C A G C A	ан а	A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T			Last Last Last Last Last Last C 28 C	Seq. End 1 933 1 933 1 933 1 933 0 933 0 933 0 933 1 933	2033 933 933 933 933 932 932 932 932 932 9

Figure 1. The new PopSet record display showing the Study Details, the Sequences list, and the submitted Alignment for a phylogenetic set (PopSet: 298351991) of apoliprotein B sequences from mammals. The Study Details shows the title of the study with a link to the citation in PubMed and the full-text in PMC (not shown) when available. The list of sequences provides the sequence titles and a link to each record in the Nucleotide database. The submitted Alignment is displayed in the embedded Graphical Sequence Viewer.

CLUSTAL, Nexus, and Phylip – are also available making the alignments easy to use for local analysis.

Improved PopSet-PopSet Connections

PopSet now features more explicit connections between PopSets associated with the same study. As always, following the link from a PubMed record retrieves all PopSets for molecules used in the study. In the previous version of PopSet, however, it was not easy to navigate from one PopSet to others that are part of the same study. The PopSet-PopSet link now provides rapid access to related PopSets. The related PopSets also are listed "Other data sets from this study" in the right-hand Discovery Column of the full record. Figure 3 shows the items in the Discovery Column and the corresponding related data in PopSet and PubMed.

	Format			Choose Destinati	ion 📔
arniv	Summary PopSet		(APOB) gene,	U	Clipboard
opSet: 29	GenBank			0	0
enBank	FASTA			Download 1 items.	
	FASTA (text)			Format	
	ASN.1			and the second se	
<u>o to:</u> ⊘	Revision History			PopSet	-
tudy D	Accession List	Apply		GenBank	
•	GI List			FASTA	
attern a.		n	of the mammalian		inferred from
ultiple r	nuclear gene se	quences.		XML	
zirik,E., M	urphy,W.J., Koepfli,	K.P., Johnson,	N.E., Dragoo, J.W., Wayr	INSDSeq.XML TinySeq.XML	3.J.
010) Mol. Pt	hylogenet. Evol. 56:(1)	49-63		Feature Table	
AID: 201382	20 Citation			FASTA plus Gap	
	and a state of the			CLUSTAL	
				Nexus	
<u>o to:</u>				Phylip	-
<u>J930905.1</u> <u>J930904.1</u> J930903.1	Ursus americanus Ailuropoda melano Ailurus fulgens apo	gi 298351991		GGAAACAATGAGAACA	IGCATTGAGGCCCACGTAGGAATAAATGGAGAAGCCAATCT
<u>J930904.1</u>	Ailuropoda melano	gi 298351991 gi 298351993		GGAACAATGAGAACA GGAACAATGAGAACA GGAAACAATGAGAACA GGAAACAATGAGAACA GGAAACAATGAGAACA GTAACAATGAGAAACA GTAACAATGAAAACA GTAACAATGAGAAACA GGAACAATGAGAACA GGAACAATGAGAACA GGAACAATGAGAACA GGAACAATGAGAACA GGAACAATGAGAACA GGAACAATGAGAACA GGAACAATGAGAACA GGAACAATGAGAACA	IGGATTGAGGCCCACGTAGGAATAAATGGAGAAGCCAATCT IGCATTGAGGCCCACGTAGGAATAAATGGAGAAGCCAATCT IGCATTGAGGCCCACGTAGGAATAAATGGAGAAGCCAATCT IGCATTGAGGCCCACGTAGGAATAAATGGAGAAGCCAATCT IGCATTGAGGCCCATGTAGGAATAAATGGAGAAGCCAATCT IGCATTGAGGCCCATGTAGGAATAAATGGAGAAGCCAATCT IGCATTGAGGCCCATGTAGGAATAAATGGAGAAGCCAATCT IGCATTGAGGCCCATGTAGGAATAAATGGAGAAGCCAATCT IGCATTGAGGCCCCATGTAGGAATAAATGGAGAAGCCAATCT IGCATTGAGGCCCCATGTAGGAATAAATGGAGAAGCCAATCT IGCATTGAGGCCCCATGTAGGAATAAATGGAGAAGCCAATCT IGCATTGAGGCCCCATGTAGGAATAAATGGAGAAGCCAATCT IGCATTGAGGCCCCACGTAGGAATAAATGGAGAAGCCAATCT IGCATTGAGGCCCCACGTAGGAATAAATGGAGAAGCCAATCT IGCTTTGAGGCCCCACGTAGGAATAAATGGAGAAGCCAATCT IGTGTTGAGGCCCCACTAGGAATAAATGGAGAAGCCAACCT IGTGTTGAGGCCCCACTAGGAATAAATGGAGAAGCCAACCT IGCTTTGAGGCCCCACATAGGAATAAATGGAGAAGCCAACCT IGCTTGAGGCCCCACATAGGAATAAATGGAGAAGCCAACCT IGCTTGAGGCCCCACATAGGAATAAGTGGAGAAGCCAACCT IGCATTGAGGCCCCACATAGGAATAAGTGGAGAAGCCAACCT IGCATTGAGGCCCCACATAGGAATAAGTGGAGAAGCCAACCT IGCATTGAGGCCCCACATAGGAATAAGTGGAGAAGCCAACCT IGCATTGAGGCCCCACATAGGAATAAGTGGAGAAGCCAACCT IGCATTGAGGCCCCACATAGGAATAAGTGGAGAAGCCAACCT IGCACTGAGGCCCCACATAGGAATAAGTGGAGAAGCCAACCT IGCACTGAGGCCCCACATAGGAATAAGTGGAGAAGCCAACCT

Figure 2. "Display Settings" (upper left) and "Send to" (upper right) menus for the new PopSet record display. PopSet retains its own separate sequence record formats (FASTA, GenBank, ASN.1). These are displayed within the PopSet database rather than in the sequence databases. Download options for PopSets with alignments include popular multiple-alignment formats such as CLUSTAL (lower panel).

Analysis Tools: BLAST and Tree View

For PopSets with fewer than 100 sequences, analysis tools are available at the top of the right- hand Discovery Column (Figure 3). These allow generating or re-generating an alignment with BLAST or, if a submitted alignment is present, displaying a distance tree (Tree View) based on the alignment. Figure 4 shows the results of the BLAST and Tree View tools for a phylogenetic study set that has a submitted alignment. The link to run BLAST is especially useful in cases where the set does not contain a submitted alignment, for example PopSet: 338197537. In such cases the Tree View can be invoked after running the BLAST alignment through the "Distance tree of results" link on the BLAST output.

	1.	Laurasiatheria RASA2 gene, par phylogenetic study, 34 aligned ser UID: 298352625		
Analyze this data set		Carnivora recombination activation	on protein 2 (RAG2) gene, partial cds.	
Run BLAST alignment	2.	phylogenetic study, 34 aligned se		
Tree View		UID: 298352554	2	
		Carnivora protein tyrosine phosp	hatase receptor type G (PTPRG) gene, partial cds.	
Article reporting this data set	3.	phylogenetic study, 39 aligned se		
Pattern and timing of diversification of the mammalian order Carnivora inferred from		UID: 298352475		
multiple nuclear gene [Mol Phylogenet Evol. 2010]		Carnivora prepronociceptin (PNC	DC) gene, partial cds.	
	4.	phylogenetic study, 35 aligned ser UID: 298352403	Mol Phylogenet Evol. 2010 Jul;56(1):49-63. Epub 2010 Feb 4.	2000) #5800
Other data sets from this study			Pattern and timing of diversification of th Carnivora inferred from multiple nuclear	
Laurasiatheria RASA2 gene, partial sequence.				
Carnivora recombination activation protein 2	5.	phylogenetic study, 36 aligned set UID: 298352366	Eizirik E, Murphy WJ, Koepfli KP, Johnson WE, Dragoo JW Faculdade de Biociências, PUCRS, Porto Alegre, RS, Brazil. ed	
(RAG2) gene, partial cds.		Carnivora growth hormone recept	Abstract	
Carnivora protein tyrosine phosphatase receptor type G (PTPRG) gene, partial cds.	6.	phylogenetic study, 23 aligned ser UID: 298352320	The mammalian order Carnivora has attracted the attention for decades, leading to intense interest in defining its supra	
Carnivora prepronociceptin (PNOC) gene, partial cds.		UID: 298352320	years, major changes to the topological structure of the car	rnivoran tree have been proposed and
Laurasiatheria PLP1 gene, partial sequence.	7.	Laurasiatheria FES gene, partial	supported by various molecular data sets, radically changin composition in this order. Although a sequence of molecular	
Carnivora growth hormone receptor (GHR) gene, partial cds.	<i>'</i> .	phylogenetic study, 25 aligned set UID: 298352295	consensus with respect to most inter-familial relationships, carnivoran lineages (both feliform and caniform) in an integ	no analysis so far has included all
Laurasiatheria FES gene, partial sequence.		Laurasiatheria FBN1 gene, partia	comparative patterns of diversification. Moreover, no study	
Laurasianena i Lo gene, partiai sequence.	8.	phylogenetic study, 44 aligned se	divergence dates among all carnivoran families, which is a	
Laurasiatheria FBN1 gene, partial sequence.		UID: 298352251	attempt to understand the patterns and tempo of diversifica have investigated the phylogenetic relationships among ca	
Laurasiatheria CHRNA1 gene, partial sequence.	9.	Laurasiatheria CHRNA1 gene, p phylogenetic study, 27 aligned se	molecular dating analyses of the inferred nodes. We asser containing 14 genes (7765 bp), most of which have not be	en previously used in supra-familial
Carnivora brain-derived neurotrophic factor (BDNF) gene, partial cds.		UID: 298352224	carnivoran phylogenetics, for 50 different genera represent this data set led to consistent and robust resolution of all si	upra-familial nodes in the carnivoran
See all		Carnivora brain-derived neurotro	tree, and allowed the construction of a molecular timescale	of or the evolution of this mammalian
000 00	10.	phylogenetic study, 42 aligned se UID: 298352143	order. Copyright 2010 Elsevier Inc. All rights reserved.	
Related information		Carnivora ATP7A (ATP7A) gene,	partial cds	
PopSet	11.			
Nucleotide	1050	UID: 298352074		
Protein				
PubMed	12.	Carnivora APP gene, partial sequences		
Taxonomy	12.	phylogenetic study, 34 aligned ser UID: 298352040	quences	
		Carnivora adenosine A3 recepto	r (ADORA3) gene, partial cds.	
	13.	phylogenetic study, 45 aligned set UID: 298351898	quences	

Figure 3. The Discovery column (left-hand image) for a PopSet record showing related PopSets (center image) and result of following the link to PubMed (right-hand image). The Discovery Column has Analysis Tools, a database ad for PubMed showing the article title, an ad for related PopSets ("Other datasets from this study"), and the traditional Links menu – now shown as Related information. Following the "See all ..." link in the related PopSets ad or the PopSet link produces the results shown in the center image, PopSet for other sequence targets reported in the linked PubMed citation. Following the linked article title in the PubMed ad or the PubMed link in the Related Information section retrieves the citation.

Summary

The NCBI PopSet database has been fully updated to the new Entrez system and includes new record displays and better access to related information. These improvements will make the growing collection of PopSets easier to access, download, and analyze.

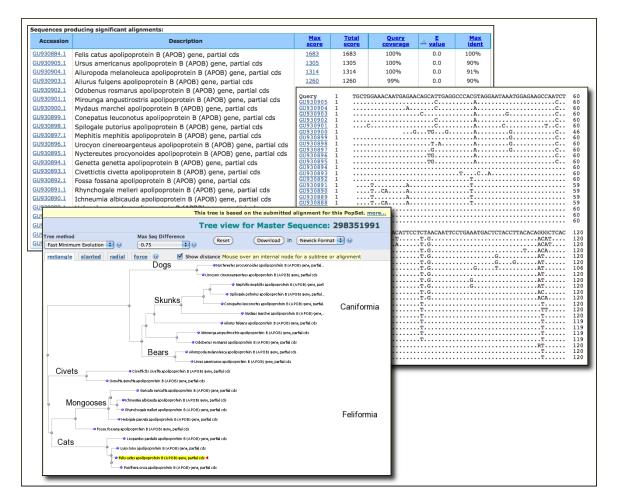


Figure 4. Results of Analysis Tools links "Run BLAST" and "Tree View" from PopSet: 298351991. The BLAST search is implemented using the first sequence as a query against the remaining members of the PopSet. The results are presented in BLAST flat query anchored format with identities shown as dots (upper images). The Tree View link uses submitted alignment in the BLAST Tree View service (lower image). In this case the result shows a molecular phylogeny for the sequences in the set, supporting major groups of the mammalian order carnivora. Names of groups were added manually and are not produced by the software.

New My NCBI Interface

My NCBI now has customizable modules making it even easier to manage your NCBI preferences, collections, bibliographies, saved searches, and more. A video highlighting the new homepage and features is on the NCBI YouTube Channel.

Search NC	BI database	8			• X	Saved Searches				
Search :	PubMed		•			Search N	lame		What's New	Last Searched
				(Search)		PopSet Searches			INOW	Soarchou
	g the "Search" se's homepage		out any terms listed in the se	arch box will transport yo	ou to	popset_pubmed[Filter] AND alignmen popset_popset[nt_present[Filter] AND	0	23	2 days ago
						PubMed Searches				
Ay Bibliog	graphy				• •	blast AND sequence analysis[Mesh] biology[mesh]	AND computational	0	Q	2 days ago
Your biblio	graphy contain	s 11 items						Manage	Saved Se	arches »
	graphy is priva									
Most rece	ently added cit	ations:				Collections				
					0	Collection Name	Items	Privac	-	Туре
			F, McGinnis SD, Pechous S, I ducation resources of the Nat			My Bibliography	O 11	Private	2	indard
Biotechn	ology Informati	on, Brief Bi	oinform. 2010 Nov;11(6):563-	9. Epub 2010 Jun 22.		Other Citations	÷ ··	Drivete	010	-dard
Review.	Publied PMID:	20570644;	PubMed Central PMCID: PI	My NCBI: No	ew Ho	me Page		_	7	Jaio
Sayers E	W, Barrett T, B	enson DA,	Bolton E, Bryant SH, Canes	Sum 48 videos	⇒ Subsc	vibe	You	Tuhe		Aed
DM, DIC	uccio M, Feder	hen S, Feol	o M, Fingerman IM, Geer LY							
	man D Linmon	DILUT		-				TUNO		Set
Miller V,	Mizrachi I, Oste	ell J, Panchi	Madden TL, Madej T, Maglot enko A, Phan L, Pruitt KD, S							iet ctions »
Miller V, Sherry S	Mizrachi I, Oste T, Shumway M	IJ, Panchi Sirotkin K,	Madden TL, Madej T, Maglot enko A, Phan L, Pruitt KD, S Slotta D, Souvorov A, Stard				104			iet ctions »
Miller V, Sherry S Wagner	Mizrachi I, Oste T, Shumway M L, Wang Y, Wilt or Biotechnolog	ell J, Panch , Sirotkin K, bur WJ, Yas y Informatic	Madden TL, Madej T, Maglot enko A, Phan L, Pruitt KD, S Slotta D, Souvorov A, Starc chenko E, Ye J. <u>Database re</u> on. Nucleic Acids Res. 2011							
Miller V, Sherry S Wagner	Mizrachi I, Oste T, Shumway M L, Wang Y, Wilt or Biotechnolog 38-51. Epub 20	ell J, Panch , Sirotkin K, bur WJ, Yas y Informatic	Madden TL, Madej T, Maglot enko A, Phan L, Pruitt KD, S Slotta D, Souvorov A, Staro chenko E, Ye J. <u>Database re</u>			Home Page				
Miller V, Sherry S Wagner I <u>Center fo</u> issue):D3	Mizrachi I, Oste T, Shumway M L, Wang Y, Wilt or Biotechnolog 38-51. Epub 20	ell J, Panch , Sirotkin K, bur WJ, Yas y Informatic	Madden TL, Madej T, Maglot enko A, Phan L, Pruitt KD, S Slotta D, Souvorov A, Starc chenko E, Ye J. <u>Database re</u> on. Nucleic Acids Res. 2011	My N	ICBI	Home Page				
Miller V, Sherry S Wagner I <u>Center fo</u> issue):D3	Mizrachi I, Oste T, Shumway M L, Wang Y, Wilt or Biotechnolog 38-51. Epub 20	ell J, Panch , Sirotkin K, bur WJ, Yas y Informatic	Madden TL, Madej T, Maglot enko A, Phan L, Pruitt KD, S Slotta D, Souvorov A, Starc chenko E, Ye J. <u>Database re</u> on. Nucleic Acids Res. 2011	My N	ICBI					
Miller V, Sherry S Wagner I <u>Center fo</u> issue):D3	Mizrachi I, Oste T, Shumway M L, Wang Y, Wilt or Biotechnolog 38-51. Epub 20	ell J, Panch , Sirotkin K, bur WJ, Yas y Informatic	Madden TL, Madej T, Magiot enko A, Phan L, Pruitt KD, S Slotta D, Souvorov A, Starc chenko E, Ye J. <u>Database ra</u> <u>m</u> , Nucleic Acids Res. 2011 PubMed PMID: 21097890; 8	My N	ICBI	Home Page				
Miller V, Sherry S Wagner I <u>Center fo</u> issue):D3	Mizrachi I, Oste T, Shumway M L, Wang Y, Wilt or Biotechnolog 38-51. Epub 20	ell J, Panch , Sirotkin K, bur WJ, Yas y Informatic	Madden TL, Madej T, Magiot enko A, Phan L, Pruitt KD, S Slotta D, Souvorov A, Starc chenko E, Ye J. <u>Database ra</u> <u>m</u> , Nucleic Acids Res. 2011 PubMed PMID: 21097890; 8	My N	ICBI	Home Page				
Miller V, Sherry S Wagner I <u>Center fo</u> issue):D PMC301	Mizrachi İ, Oste T, Shumway M L, Wang Y, Wills z Biotechnolog 38-51. Epub 20 3733.	ell J, Panch , Sirotkin K, bur WJ, Yas y Informatic	Madden TL, Madej T, Magiot enko A, Phan L, Pruitt KD, S Slotta D, Souvorov A, Starc chenko E, Ye J. <u>Database ra</u> <u>m</u> , Nucleic Acids Res. 2011 PubMed PMID: 21097890; 8	My N	ICBI	Home Page				
Miller V, Sherry S Wagner I <u>Center fo</u> issue):D PMC301	Mizrachi İ, Oste T, Shumway M L, Wang Y, Wills z Biotechnolog 38-51. Epub 20 3733.	ell J, Panchi , Sirotkin K, bur WJ, Yas 10 Nov 21.	Madden TL, Madej T, Magiot enko A, Phan L, Pruitt KD, S Slotta D, Souvorov A, Starc chenko E, Ye J. <u>Database ra</u> <u>m</u> , Nucleic Acids Res. 2011 PubMed PMID: 21097890; 8	My N	ICBI	Home Page				
Miller V, Sherry S Wagner / Center fc issue):D: PMC301	Mizrachi I, Oste T, Shumway M L, Wang Y, Wit J Biotechnolog 38-51. Epub 20 3733.	ell J, Panch , Sirotkin K, bur WJ, Yas y Informatic	Madden TL, Madej T, Magiot enko A, Phan L, Pruitt KD, S Slotta D, Souvorov A, Starc chenko E, Ye J. <u>Database ra</u> <u>m</u> , Nucleic Acids Res. 2011 PubMed PMID: 21097890; 8	My N	ICBI	Home Page				• 2
Miller V, Sherry S Wagner 1 <u>Center fc</u> issue):D: PMC301 Recent Ac	Mizrachi I, Oste T, Shumway M L, Wang Y, Wilt zr Biotechnolog 38-51. Epub 20 3733.	ell J, Panchu , Sirotkin K, bur WJ, Yas y Informatic 110 Nov 21.	Madden TL, Madej T, Magiot enko A, Phan L, Pruit KO, S Slotta D, Souvorov A, Starc chenko E, Ye J, Databaser M, Jana M, Jana M, Sang M, Nucleic Acids Res. 2011 PubMed PMID: 21097890; f see all citations ≥	My N	ICBI	Home Page	NC	R		
Miller V, Sherry S Wagner I Issue):D: PMC301 Recent Ac	Mizrachi I, Oste T, Shumway M L, Wang Y, Wilt Preletechnolog 3733. tivity Database Protein	II J. Panchi, Sirotkin K, Jur WJ. Yasa y Informatic 10 Nov 21.	kadden TL, Madej T, Magiot enko A, Phan L, Pruit KO, S Slotta D, Souvorov A, Starc chenko E, Ye J, Databaser M In, Nucleic Acids Res. 2011 PubMed PMID: 21097690; f see all citations ≥ zebrafishiOrganism] AND	My N	ICBI	Home Page	NC	B		
Miller V, Sherry S Wagner I Issue):D: PMC301 Recent Ac	Mizrachi I, Oste T, Shumway M L, Wang Y, Wilt Preletechnolog 3733. tivity Database Protein	II J. Panchi, Sirotkin K, Jur WJ. Yasa y Informatic 10 Nov 21.	kadden TL, Madej T, Magiot enko A, Phan L, Pruit KO, S Slotta D, Souvorov A, Starc chenko E, Ye J, Databaser M In, Nucleic Acids Res. 2011 PubMed PMID: 21097690; f see all citations ≥ zebrafishiOrganism] AND	My N	ICBI	Home Page	NC	B	I	
Miller V, Sherry S Wagner I Issue):D: PMC301 Recent Ac	Mizrachi I, Oste T, Shumway M L, Wang Y, Wilt Preletechnolog 3733. tivity Database Protein	II J. Panchi, Sirotkin K, Jur WJ. Yasa y Informatic 10 Nov 21.	kadden TL, Madej T, Magiot enko A, Phan L, Pruit KO, S Slotta D, Souvorov A, Starc chenko E, Ye J, Databaser M In, Nucleic Acids Res. 2011 PubMed PMID: 21097690; f see all citations ≥ zebrafishiOrganism] AND	My N	ICBI	Home Page	NC	B	I	
Miller V, Sherry S Wagner I <u>Center fr</u> issue):D; PMC301 Recent Ac	Mizrachi I, Oste T, Shumway M L, Wang Y, Wilt Preletechnolog 3733. tivity Database Protein	II J. Panchi, Sirotkin K, Jur WJ. Yasa y Informatic 10 Nov 21.	kadden TL, Madej T, Magiot enko A, Phan L, Pruit KO, S Slotta D, Souvorov A, Starc chenko E, Ye J, Databaser M In, Nucleic Acids Res. 2011 PubMed PMID: 21097690; f see all citations ≥ zebrafishiOrganism] AND	My N Tour the r	ICBI new My	Home Page	NC.	B		• 2
Miller V, Sherry S Wagner I <u>Center fr</u> issue):D; PMC301 Recent Ac	Mizrachi I, Oste T, Shumway M L, Wang Y, Wilt Preletechnolog 3733. tivity Database Protein	II J. Panchi, Sirotkin K, Jur WJ. Yasa y Informatic 10 Nov 21.	kadden TL, Madej T, Magiot enko A, Phan L, Pruit KO, S Slotta D, Souvorov A, Starc chenko E, Ye J, Databaser M In, Nucleic Acids Res. 2011 PubMed PMID: 21097690; f see all citations ≥ zebrafishiOrganism] AND	My N Tour the r	Add to	Home Page	NC.	B		

Transcriptome Shotgun Assembly (TSA) Database Available for BLAST

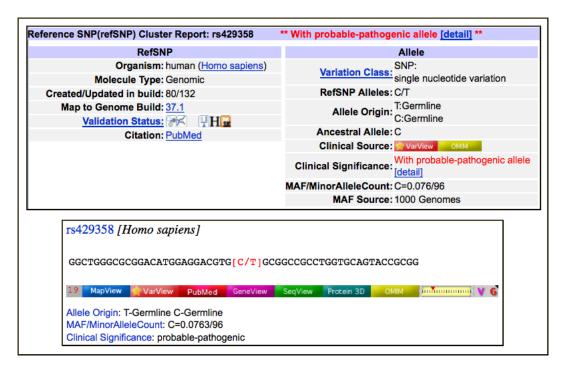
The Transcriptome Shotgun Assembly (TSA) BLAST database is now available from the database list for the main NCBI BLAST services. TSA is an archive of computationally assembled mRNA sequences from primary data such as Expressed Sequence Tag (EST) and raw sequence reads. These sequences were previously a part of the BLAST nucleotide nr (nt) database but have been moved because of their increasing numbers and special characteristics. The TSA page has more information on the nature and sources of TSA sequences.

New Attributes for Human Variants in dbSNP

New attributes related to allele origin, clinical significance, and population genetics are available in dbSNP. These attributes allow searching and filtering of human variations for the characteristics listed below.

- 1. Allele Origin: Summarizes the reported origin(s) of the variant allele asserted by each submitter for the submitted SNP (ss). Current values are germline, somatic, and unknown. Additional attributes will be added in the future including not-tested, tested-inconclusive, and other.
- 2. Clinical significance: Reports potential health impact of the allele. Possible values:
 - unknown
 - untested
 - non-pathogenic
 - probable-non-pathogenic
 - probable-pathogenic
 - pathogenic
 - drug response
 - histocompatibility
 - other
- 3. **Global minor allele frequency (MAF):** Shows the minor allele frequency for each RefSNP included in a default global population. Since this is being provided to distinguish common polymorphism from rare variants, the MAF is actually the second most frequent allele value. For example, if there are 3 alleles with frequencies of 0.50, 0.49, and 0.01, the MAF will be reported as 0.49. The current default global population is 1000Genome phase 1 genotype data from 629 worldwide individuals, released in the 08-04-2010 dataset.
- 4. **Suspect:** Variation suspected to be false positive due to various artifacts.

These new attributes are shown in the images below for the rs429358 Cluster Report and Document Summary.



Please see the online help for more information and more examples.

Updated BLAST Genome Search Pages

The genome-specific BLAST pages linked to the top of the NCBI BLAST homepage and accessible from the Map Viewer homepage now use the standard BLAST form with genome specific databases. This change eliminates the older separate interface and provides the full functionality of the standard BLAST interface including the ability to adjust all algorithm parameters, the capability to edit and re-submit searches, to sort descriptions and alignments in the output, and the full range of formatting and downloading options.

NLM Contest: Show off your Apps! Invitation to Submit Applications that Work with NLM Biomedical Data

The National Library of Medicine (NLM) is challenging people to create innovative software applications that use the Library's vast collection of biomedical data. The purpose of this contest is to foster the development of innovative software applications that will further NLM's mission of aiding the dissemination and exchange of scientific and other information pertinent to medicine and public health. Winners will be recognized at an awards ceremony at the National Library of Medicine and links to their application will be publicized on NLM Web sites. The NLM "Show Off Your Apps" Challenge is open to individuals over the age of 18, teams of individuals, and organizations in the United States. Eligible software applications must make use of NLM's vast collection of biomedical data including downloadable data sets, application programming interfaces, and/or software tools. The challenge.gov website has detailed information on the contest.

Applications should be submitted to the challenge.gov site by August 31, 2011.

New Videos on NCBI's YouTube Channel

In addition to the video introducing the new My NCBI mentioned above, four other instructional videos recently became available on NCBI's YouTube channel:

- Saving search results in My NCBI Collections
- Loading sequences and adjusting graphical views in NCBI's Genome Workbench
- Requesting permission to use controlled access data in dbGaP
- Using BLAST from NCBI's graphical sequence viewer

The Sequence Read and Trace Archive Databases to Continue

Recently, NCBI announced that the Sequence Read Archive (SRA) and Trace Archive repositories would be discontinued due to budget constraints (NCBI News, March 2011). However, with the commitment of interim funding and a plan for future support developed in collaboration with other NIH Institutes and NIH grantees, NCBI will now continue to accept submissions and maintain the Sequence Read Archive (SRA) and Trace Archive repositories for high-throughput sequence data. These repositories will now focus on high-throughput data that support other kinds of data at the NCBI including:

- RNA-Seq, ChIP-Seq, and epigenomic data that are submitted to GEO
- Genomic and Transcriptomic assemblies that are submitted to GenBank
- 16S ribosomal RNA data associated with metagenomics that are submitted to GenBank

The full announcement on the NCBI site has more details.

BLAST 2.2.25+ Release and New Set-up Instructions

Stand-alone BLAST+ (v2.2.25) is now on the FTP site. Improvements include hardmasking of databases, faster formatting of databases using makeblastdb, XML and best hit options for Blast2Sequences, multiple query psiblast, selection of any master sequence in psiblast with multiple alignment input, and query and subject length in tabular output. The BLAST News has more detailed information on changes. Detailed set-up instructions for standalone BLAST are now a part of the BLAST User Manual on the NCBI Bookshelf.

Microbial Genomes Update

One hundred thirty five finished microbial genomes were released between March 1 and May 31, 2011. The original sequence data files submitted to GenBank/EMBL/DDBJ are available in the Bacteria directory in the /genbank/genomes area of the GenBank FTP site. One hundred twelve RefSeq provisional versions were made from a selected set of finished genomes. These are available from the /genomes/Bacteria directory on the FTP site.

In addition, 305 microbial whole genome shotgun-sequencing projects were added to GenBank during this period. The original submitted files are available in the Bacteria_DRAFT directory in the GenBank genomes area. RefSeq provisional versions of 84 of these projects are available in the /genomes/Bacteria_DRAFT area of the FTP site.

All GenBank and RefSeq microbial genomes are incorporated in the NCBI integrated Entrez search and retrieval system and the BLAST sequence similarity search service.

RefSeq News

RefSeq Release 47 is available through Entrez, BLAST, and the RefSeq FTP site. The current release includes 17.6 million sequence records from 12,000 organisms. Release notes provide more detailed information.

GenBank News

GenBank release 183 is available through the NCBI web and FTP sites. The current release incorporates data available as of Apr 11, 2011 and, with the whole-genome shotgun portion, contains 317,952,894,329 bases from 198,156,212 sequence records. Release notes describe the current state of data and upcoming changes.

NCBI Discovery Workshops at Washington University: July 26-27, 2011

NCBI will present a two-day workshop on July 26 and 27th, at Washington University in St. Louis, Missouri. The course is free and is open to anyone interested in NCBI resources. The workshops provide hands-on experience exploring practical examples using tools and databases on the NCBI website. The four workshops are Sequences, Genomes, and Maps; Proteins, Domains and Structures; NCBI BLAST Services; and Human Variation and Disease Genes. The Discovery Workshops page has more information.

Announce Lists and RSS Feeds

Eighteen topic-specific mailing lists are available that 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. Subscribe to the NCBI Announce list to receive updates on the NCBI News.

Twelve RSS feeds are now available from NCBI including news on PubMed, PubMed Central, NCBI Bookshelf, LinkOut, HomoloGene, UniGene, and NCBI Announce.

NCBI's Facebook page and Twitter feed also provide updates on NCBI resources.

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.