

Contract No. HHSN266200400006C

Immune Epitope Database and Analysis Program

2006 Annual IEDB Compendium

La Jolla Institute for Allergy and Immunology
9420 Athena Circle
La Jolla, CA 92037

858-752-6923
858-752-6987 (fax)
wfleri@liai.org

January 2007
Revised February 2007

Table of Contents

Table of Contents	i
Introduction	1
1 Antibody and T Cell Epitopes	2
2 Website Features	22
2.1 Home Page	22
2.2 Query	23
2.2.1 Browse for Records	23
2.2.1.1 Browse Records by MHC Allele	23
2.2.1.2 Browse Records by Source Species	24
2.2.1.3 Browse Records by 3D Structure	24
2.2.2 Perform a Search	24
2.2.3 Perform a Simple Query	24
2.2.4 Search and Simple Query – Results Summary Page	26
2.2.5 Advanced Query	28
2.2.5.1 View Query Criteria	28
2.2.5.2 Save Queries	28
2.2.5.3 View Advanced Query Results	29
2.2.5.4 Download Advanced Query Results	29
2.2.5.5 Accessing the EpitopeViewer	29
2.2.6 Finders Overview	30
2.2.6.1 MHC Allele Finder	30
2.2.6.2 Assay Type Finder	30
2.2.6.3 Disease Finder	31
2.2.6.4 Source Finder	32
2.2.6.5 Species Finder	33
2.2.7 Analysis Tools	34
2.3 Resources	34
2.3.1 Analysis Resource	34
2.3.1.1 Analysis Tools	35
2.3.1.1.1 Population coverage	35
2.3.1.1.2 Epitope conservancy	36
2.3.1.1.3 Visualization Tool	37
2.3.1.1.4 Tool for mapping linear epitopes to the PDB structures	37
2.3.1.2 Predictive tools	38
2.3.1.2.1 MHC binding prediction methods	38
2.3.1.2.1.1 Artificial Neural Network	38
2.3.1.2.1.2 Average Relative Binding (ARB)	38
2.3.1.2.1.3 Stabilized Matrix Method (SMM)	38
2.3.1.2.1.4 MHC Class I Binding Prediction Resource	38
2.3.1.2.2 Proteasomal cleavage and TAP transport predictions	39
2.3.1.2.3 Sequence based antibody epitope prediction	39
2.3.2 Forums	40
2.3.3 Database Export	40

2.3.4	Downloads	41
2.3.5	View External Links	41
2.4	Account Information	41
2.4.1	Become a Registered User	41
2.4.2	Registered User Login	41
2.4.3	Modify Account Information	42
2.4.4	Terms of Use	42
2.5	Support Overview	42
2.5.1	Submit Feedback	42
2.5.2	Submit Help Request	42
2.5.3	View Online Help	43
2.6	About IEDB	43
2.6.1	Tour the IEDB	43
2.6.2	Acknowledgements	44
2.6.3	Publications	44
2.6.4	Citing the IEDB	44
3	Scientific Publications	46
4	References	47

Introduction

The Immune Epitope Database and Analysis Resource (IEDB) is a public repository of immune epitope data and is sponsored by the National Institute for Allergy and Infectious Diseases (NIAID). The IEDB development started in December 2003 and it became available to the public in a beta test phase on 15 February 2006. The IEDB contains data related to antibody and T cell epitopes for humans, non-human primates, rodents, and other animal species. The IEDB also makes available a variety of analytical and epitope prediction tools and resources within its Analysis Resource.

This second Annual Compendium of the Immune Epitope Database and Analysis Resource consists of three sections. The first section contains a list of the antibody and T cell epitope information in the database as of 11 January 2007. The second section lists and describes the various features of the IEDB website available by the end of 2006. The third section lists the scientific publications in 2006 for which the IEDB played a contributory role.

Since the publication of last year's 2005 Annual Compendium, the quantity of data available in the IEDB has increased significantly with the addition of 1422 fully curated references. The Category A-C priority pathogen literature and practically all of the emerging/re-emerging infectious disease references (<http://www3.niaid.nih.gov/research/topics/emerging/list.htm>) were processed and curated by the end of 2006. The curation of malaria references was started, and a large quantity of MHC binding data was submitted and incorporated.

Two new website features are described in this Compendium – Browse Records by 3D Structure and Search using a URL. A number of other new website features and capabilities were still in the test phase at the end of 2006, including three antibody prediction tools (BepiPred, DiscoTope, and ElliPro), and an improved epitope conservancy tool. These will be introduced in 2007 and will be fully described in the 2007 Compendium.

1 Antibody and T Cell Epitopes

Many new references and many new pathogens were added to the IEDB in 2006, as demonstrated in Table 1. The table lists the number of B cell and T cell epitopes in the database by source species as of 31 December 2005 and 11 January 2007. Last year's compendium listed epitopes for 401 source species. In 2006, 336 new source species were added to the list and most of the previously listed species gained additional epitopes. It should be noted that the source species is the species from which the epitopes originate, and may not be from an infecting organism or vaccine target. The curation of MHC binding peptides, cross-reactive epitopes, and autoimmune epitopes has resulted in the appearance of human and mouse epitopes on the list. Although the curation of autoimmune epitopes has not been a priority to date, they have entered the IEDB as a result of importing large epitope datasets and of curating them when they appear in priority infectious disease references.

The NCBI taxonomy ID is provided for each source species and strain. This list differs from the one available on the **Browse Records by Source Species** (<http://immuneepitope.org/browse.do?dispatch=loadSpeciesBrowser>) web page because the table below only counts distinct epitopes, not records. A **distinct epitope** is an epitope with given characteristics, regardless of the reference, that has a positive qualitative binding value for antigen or MHC association. The **Browse Records by Source Species** (<http://immuneepitope.org/browse.do?dispatch=loadAlleleBrowser>) list provides a count of records, which includes molecular structures in the database that have positive or negative qualitative binding values. A **record** contains data about one epitope (or structure) in one reference, where a reference is a scientific article or a direct data submission from a researcher. The IEDB collects both positive and negative data, a helpful feature for epitope prediction tool developers. The reader can obtain more information about the epitopes of a specific source species by going to <http://www.immuneEpitope.org> and performing a query.

In the table below, the leftmost column labeled "New" indicates with an "X" if epitopes for the species/strain was added to the IEDB in 2006. A font color of red is used to highlight the information in the row. The columns labeled "B-05", "T-05", "B-06", and "T-06" indicate the cumulative number of distinct B cell and T cell epitopes in the database at the end of 2005 and 2006, respectively. The two rightmost columns display the differences in the B and T cell epitope counts from 2005 to 2006. In a few cases the count is negative, which indicates that the source species of the 2005 epitope was reassigned during a re-curation process. The changes in B and T cell epitope counts are shown in red. In 2006, the number of distinct B cell epitopes increased by 4102, from 1154 to 5256, and the number of distinct T cell epitopes increased by 11706, from 4626 to 16332.

Table 1 Summary of B and T cell epitopes contained in the IEDB

New 2006	NCBI TAX ID	SPECIES/STRAIN	B-05	T-05	B-06	T-06	Δ B	Δ T
X	5755	<i>Acanthamoeba castellanii</i>			14	1	14	1
X	714	<i>Actinobacillus actinomycescomitans</i>				4		4
	228399	<i>Actinobacillus pleuropneumoniae</i> serovar 1 str. 4074	2		2			
X	10508	<i>Adenoviridae</i>			3	2	3	2
	1049	<i>Allochromatium vinosum</i>		1		1		
	5599	<i>Alternaria alternata</i>	3		3			
	4212	<i>Ambrosia artemisiifolia</i>		1				-1
X	4215	<i>Ambrosia elatior</i>				1		1
	171929	<i>Anacardium occidentale</i>	11		11			
	770	<i>Anaplasma marginale</i>	14	6	14	16		10
X	320483	<i>Anaplasma marginale</i> str. Florida				15		15
	234826	<i>Anaplasma marginale</i> str. St. Maries		1		1		
	46607	Andes virus	50		50			
X	3702	<i>Arabidopsis thaliana</i>				1		1
	3818	<i>Arachis hypogaea</i>	39	8	39	8		
	5085	<i>Aspergillus fumigatus</i>	19		19			
	5064	<i>Aspergillus restrictus</i>	13		13			
	7604	<i>Asterias rubens</i>		1		1		
X	172851	<i>Avian hepatitis E virus</i>			7		7	
	31563	Avian rotavirus A	4		4			
	5865	<i>Babesia bovis</i>		10		10		
	5868	<i>Babesia microti</i>	2		2			
	1392	<i>Bacillus anthracis</i>	2	5	24	7	22	2
X	1396	<i>Bacillus cereus</i>				2		2
X	1402	<i>Bacillus licheniformis</i>				5		5
	37962	Bayou virus	1		1			
X	31715	<i>Bean-pod mottle virus</i> (strain Kentucky G7)				1		1
X	323422	<i>beta proteobacterium</i> MWH-C4			13	30	13	30
X	3505	<i>Betula pendula</i>			1		1	
	520	<i>Bordetella pertussis</i>		3	2	3	2	
	12455	Borna disease virus		1		1		
	29518	<i>Borrelia afzelii</i>		1		1		
	139	<i>Borrelia burgdorferi</i>		13	43	60	43	47
X	224326	<i>Borrelia burgdorferi</i> B31			34		34	
X	64895	<i>Borrelia burgdorferi</i> group			3		3	
X	40834	<i>Borrelia duttonii</i>			7		7	
X	29519	<i>Borrelia garinii</i>			3		3	
	9913	<i>Bos taurus</i>	19	31	32	48	13	17
	11303	Bovine ephemeral fever virus	2		2			
X	11215	<i>Bovine parainfluenza virus 3</i>				1		1
	11099	Bovine viral diarrhea virus 1	1		1			
	3707	<i>Brassica juncea</i>	9		9			

New 2006	NCBI TAX ID	SPECIES/STRAIN	B-05	T-05	B-06	T-06	Δ B	Δ T
X	35802	<i>Brucella abortus</i> 1			2		2	
	29459	<i>Brucella melitensis</i>	1		4	16	3	16
	224914	<i>Brucella melitensis</i> 16M	1		6		5	
X	235	<i>Brucella melitensis</i> biovar Abortus			3	3	3	3
	236	<i>Brucella melitensis</i> biovar Ovis	2		2			
X	29461	<i>Brucella melitensis</i> biovar Suis				2		2
X	89462	<i>Bubalus bubalis</i>				1		1
X	13373	<i>Burkholderia mallei</i>				1		1
X	243160	<i>Burkholderia mallei</i> ATCC 23344				1		1
	6239	<i>Caenorhabditis elegans</i>		1		6		5
X	32019	<i>Campylobacter fetus</i> subsp. fetus			3		3	
X	197	<i>Campylobacter jejuni</i>				5		5
	5476	<i>Candida albicans</i>	2	15	2	31		16
	74724	Canine calicivirus	2		2			
X	10788	Canine parvovirus			1		1	
X	9615	<i>Canis familiaris</i>				2		2
X	11662	Caprine arthritis encephalitis virus (STRAIN G63)			6		6	
X	47929	<i>Cercopithecine herpesvirus</i> 8				66		66
	9534	<i>Cercopithecus aethiops</i>		1		7		6
X	9863	<i>Cervus nippon</i>				1		1
	7155	<i>Chironomus thummi thummi</i>	3		3			
	810	Chlamydia		1		1		
X	83560	<i>Chlamydia muridarum</i>				3		3
	813	<i>Chlamydia trachomatis</i>	10	8	11	12	1	4
X	83555	<i>Chlamydomydia abortus</i>			14		14	
	83558	<i>Chlamydomydia pneumoniae</i>		8		8		
	182082	<i>Chlamydomydia pneumoniae</i> TW-183		9		9		
X	83554	<i>Chlamydomydia psittaci</i>			64		64	
	11096	Classical swine fever virus	1		1			
X	214432	Cloning vector pscFvCA-E8VHd				1		1
	1491	<i>Clostridium botulinum</i>		6	120	21	120	15
	36826	<i>Clostridium botulinum</i> A	44		44			
	1502	<i>Clostridium perfringens</i>	1		2		1	
X	195102	<i>Clostridium perfringens</i> str. 13				1		1
	1513	<i>Clostridium tetani</i>		22		25		3
X	5501	<i>Coccidioides immitis</i>				1		1
X	199306	<i>Coccidioides posadasii</i>				2		2
X	9014	<i>Colinus virginianus</i>				2		2
	8932	<i>Columba livia</i>		7		7		
X	5347	<i>Coprinellus congregatus</i>				1		1
	1717	<i>Corynebacterium diphtheriae</i>		1	1	3	1	2
X	152794	<i>Corynebacterium efficiens</i>				1		1
X	1718	<i>Corynebacterium glutamicum</i>				17		17
X	10243	Cowpox virus				1		1

New 2006	NCBI TAX ID	SPECIES/STRAIN	B-05	T-05	B-06	T-06	Δ B	Δ T
X	265872	Cowpox virus (Brighton Red)				1		1
X	777	Coxiella burnetii				8		8
X	103906	Coxsackievirus B4 (strain JVB / Benschoten / New York/51)				73		73
X	10029	Cricetulus griseus				3		3
	10034	Cricetus cricetus		1		1		
	3369	Cryptomeria japonica	4	1	4	2		1
X	5807	Cryptosporidium parvum				4		4
	10358	Cytomegalovirus		23		32		9
X	7955	Danio rerio				1		1
X	243164	Dehalococcoides ethenogenes 195				1		1
	11052	Dengue virus group	1		1			
	11053	Dengue virus type 1	1		2	11	1	11
X	33741	Dengue virus type 1 (strain Singapore S275/90)				2		2
	11060	Dengue virus type 2	8	2	42	38	34	36
X	11062	Dengue virus type 2 (isolate Malaysia M2)			1		1	
	11065	Dengue virus type 2 (NGC-prototype)	10	8	12	10	2	2
X	31634	Dengue virus type 2 (strain 16681)				50		50
	31635	Dengue virus type 2 (strain 16681-PDK53)	1	2	1	2		
	11064	Dengue virus type 2 (strain Jamaica)	4		9	10	5	10
	11066	Dengue virus type 2 (strain PR159/S1)	2	2	7	2	5	
	11068	Dengue virus type 2 (strain PUO-218)	1	1	1	1		
	11069	Dengue virus type 3		9		45		36
	11070	Dengue virus type 4		10		26		16
	6954	Dermatophagoides farinae		3		3		
	6956	Dermatophagoides pteronyssinus		6	1	7	1	1
	7227	Drosophila melanogaster		1	2	3	2	2
X	11021	Eastern equine encephalitis virus			8		8	
	11268	Ebola virus		1		1		
	129003	Ebola virus strain Reston	2		2			
X	128949	Ebola virus strain Sudan Maleo-79				1		1
	128952	Ebola virus strain Zaire Mayinga	9	3	9	4		1
	31705	Echovirus 11 (strain Gregory)		1		1		
X	779	Ehrlichia ruminantium			3		3	
	12104	Encephalomyocarditis virus						
	5759	Entamoeba histolytica	4		12		8	
	10710	Enterobacteria phage lambda		7		7		
X	10665	Enterobacteria phage T4			5	20	5	20
X	11047	Equine arteritis virus			2		2	
	11665	Equine infectious anemia virus		1		1		
X	11670	Equine infectious anemia virus (CLONE 1369)				1		1
	11671	Equine infectious anemia virus (STRAIN WSU5)		6		6		
X	9796	Equus caballus			1	1	1	1

New 2006	NCBI TAX ID	SPECIES/STRAIN	B-05	T-05	B-06	T-06	Δ B	Δ T
	562	Escherichia coli		3	93	53	93	50
	83333	Escherichia coli K12		7		7		
	83334	Escherichia coli O157:H7	1		1	16		16
X	217992	Escherichia coli O6				6		6
X	57482	European bat lyssavirus 1				1		1
	11978	Feline calicivirus	4		4			
X	11980	Feline calicivirus (STRAIN JAPANESE F4)			6		6	
X	11673	Feline immunodeficiency virus				28		28
	9685	Felis catus	3	2	3	3		1
X	127586	Filamentous phage display vector f88-4			21		21	
X	127587	Filamentous phage display vector fUSE5			7		7	
	12110	Foot-and-mouth disease virus	1		1			
	12118	Foot-and-mouth disease virus - type O	1		1			
	73482	Foot-and-mouth disease virus (strain O1)	1		1			
X	31621	Four Corners hantavirus			2		2	
	263	Francisella tularensis		3		3		
	11795	Friend murine leukemia virus		2		2		
	11797	Friend murine leukemia virus (ISOLATE FB29)		1		1		
	9031	Gallus gallus	9	22	9	38		16
X	54290	GB virus C			5		5	
X	5741	Giardia intestinalis				3		3
X	184922	Giardia lamblia ATCC 50803				41		41
	3847	Glycine max	17		17			
X	9595	Gorilla gorilla gorilla				4		4
X	45219	Guanarito virus				9		9
X	114727	H1N1 subtype			1		1	
	727	Haemophilus influenzae		1		31		30
X	11599	Hantaan virus			2		2	
	11602	Hantaan virus 76-118	3	2	8	5	5	3
X	13557	Hapalemur griseus				1		1
X	32025	Helicobacter hepaticus				1		1
	210	Helicobacter pylori	1		21		20	
X	85962	Helicobacter pylori 26695			13		13	
X	63330	Hendra virus			6		6	
X	11102	Hepacivirus				3		3
	12092	Hepatitis A virus	16		16	8		8
	12098	Hepatitis A virus (STRAIN HM-175)	82		87	2	5	2
X	12099	Hepatitis A virus (STRAIN LA)				1		1
	10407	Hepatitis B virus	1	67	4	95	3	28
	10411	Hepatitis B virus (STRAIN ALPHA1)		10		19		9
	10414	Hepatitis B virus (STRAIN LSH / CHIMPANZEE ISOLATE)		1		2		1
	31512	Hepatitis B virus (subtype ADR / mutant)		1		1		
	10409	Hepatitis B virus (SUBTYPE ADR4)		10		10		

New 2006	NCBI TAX ID	SPECIES/STRAIN	B-05	T-05	B-06	T-06	Δ B	Δ T
	10410	Hepatitis B virus (SUBTYPE ADW / STRAIN 991)		4		4		
	10412	Hepatitis B virus (SUBTYPE ADW / STRAIN INDONESIA/PIDW420)		1		1		
	10413	Hepatitis B virus (SUBTYPE ADW / STRAIN JAPAN/PJDW233)		20		20		
	10415	Hepatitis B virus (SUBTYPE ADW / STRAIN OKINAWA/PODW282)		10		10		
	31515	Hepatitis B virus (subtype ADW2 variant SF)		2		2		
	10408	Hepatitis B virus (SUBTYPE ADW2)		33		33		
	45410	Hepatitis b virus (subtype ADW4 / strain brazil / isolate w4b)		2		12		10
X	31511	Hepatitis B virus subtype AD				1		1
	106821	Hepatitis B virus subtype ADW		2		11		9
	10419	Hepatitis B virus subtype ADYW		6		21		15
	10418	Hepatitis B virus subtype AYW		68		72		4
	11103	Hepatitis C virus		9	912	814	912	805
	11104	Hepatitis C virus (isolate 1)		20	37	221	37	201
X	356391	Hepatitis C virus (isolate 6a33)				5		5
X	356413	Hepatitis C virus (isolate BEBE1)				10		10
	11105	Hepatitis C virus (isolate BK)		39	180	51	180	12
X	333284	Hepatitis C virus (isolate Con1)			3	5	3	5
X	356419	Hepatitis C virus (isolate EUH1480)				4		4
X	329389	Hepatitis C virus (isolate Glasgow)				1		1
	11108	Hepatitis C virus (isolate H)		17	121	97	121	80
	63746	Hepatitis C virus (isolate H77)	1		12	63	11	63
X	356410	Hepatitis C virus (isolate HC-G9)			2	1	2	1
	11111	Hepatitis C virus (isolate HC-J2)		1				-1
	11112	Hepatitis C virus (isolate HC-J5)		1		1		
	11113	Hepatitis C virus (isolate HC-J6)		1	3	2	3	1
	11115	Hepatitis C virus (isolate HC-J8)		3	2	3	2	
	11110	Hepatitis C virus (isolate HCT18)		1	1		1	-1
X	356416	Hepatitis C virus (isolate HCV-K3a/650)				14		14
X	31644	Hepatitis C virus (isolate HCV-KF)			1	1	1	1
X	356386	Hepatitis C virus (isolate India)				1		1
	11116	Hepatitis C virus (isolate Japanese)		21	2	36	2	15
X	356411	Hepatitis C virus (isolate JFH-1)				1		1
X	356417	Hepatitis C virus (isolate JK049)				5		5
X	356415	Hepatitis C virus (isolate NZL1)			1	1	1	1
	31645	Hepatitis C virus (isolate Taiwan)		17	7	27	7	10
	11117	Hepatitis C virus (isolate TH)		3				-3
X	356421	Hepatitis C virus (isolate Th580)				4		4
X	357355	Hepatitis C virus (isolate Tr Kj)				3		3
X	356424	Hepatitis C virus (isolate VN004)				1		1
	41856	Hepatitis C virus genotype 1		18	1	24	1	6
X	40271	Hepatitis C virus genotype 2			31		31	

New 2006	NCBI TAX ID	SPECIES/STRAIN	B-05	T-05	B-06	T-06	Δ B	Δ T
X	356114	Hepatitis C virus genotype 3				3		3
X	31646	Hepatitis C virus subtype 1a			1	33	1	33
	31647	Hepatitis C virus subtype 1b	1	11	5	59	4	48
X	31649	Hepatitis C virus subtype 2a				1		1
X	31650	Hepatitis C virus subtype 2b				3		3
X	356426	Hepatitis C virus subtype 3a				2		2
X	31653	Hepatitis C virus subtype 4a			1	1	1	1
X	31654	Hepatitis C virus subtype 5a			1	1	1	1
	12475	Hepatitis delta virus		5		5		
	10426	Hepatitis delta virus (ISOLATE NAURU)		3		3		
	31764	Hepatitis delta virus (isolate woodchuck)		1		1		
X	12461	Hepatitis E virus			129		129	
X	31767	Hepatitis E virus (strain Burma)			160		160	
X	31768	Hepatitis E virus (strain Mexico)			31		31	
X	28300	Heron hepatitis B virus				1		1
	10299	Herpes simplex virus (type 1 / strain 17)						
	10301	Herpes simplex virus (type 1 / strain Angelotti)		2		2		
X	10304	Herpes simplex virus (type 1 / strain F)				1		1
	3981	Hevea brasiliensis	34		34			
X	11583	HoJo virus				1		1
	9606	Homo sapiens	27	1388	97	1966	70	578
X	129951	Human adenovirus C				1		1
	10515	Human adenovirus type 2		1		1		
	28284	Human adenovirus type 40		1		1		
	28285	Human adenovirus type 5		1		1		
	10519	Human adenovirus type 7		1		1		
	11137	Human coronavirus 229E		2		2		
X	12072	Human coxsackievirus B3			1	2	1	2
X	208726	Human hepatitis A virus			3		3	
	10298	Human herpesvirus 1		1	1	2	1	1
	10306	Human herpesvirus 1 strain KOS		2		2		
	10310	Human herpesvirus 2		6		10		4
	10313	Human herpesvirus 2 strain 333		1		1		
	10315	Human herpesvirus 2 strain HG52		4		4		
	10376	Human herpesvirus 4		10	1	26	1	16
	10377	Human herpesvirus 4 (strain B95-8)		60	1	83	1	23
X	31525	Human herpesvirus 4 (strain CAO)				1		1
	10359	Human herpesvirus 5		6		8		2
	10360	Human herpesvirus 5 strain AD169		6		18		12
	10363	Human herpesvirus 5 strain Towne		4		4		
X	10368	Human herpesvirus 6			3	1	3	1
	10370	Human herpesvirus 6 (strain Uganda-1102)		1	1	1	1	
X	32604	Human herpesvirus 6B			1	1	1	1
	57278	Human herpesvirus 7 strain JI		1		1		

New 2006	NCBI TAX ID	SPECIES/STRAIN	B-05	T-05	B-06	T-06	Δ B	Δ T
X	37296	Human herpesvirus 8			26	124	26	124
	12721	Human immunodeficiency virus		9		9		
	11676	Human immunodeficiency virus 1		15	1	77	1	62
X	11709	Human immunodeficiency virus 2				4		4
	11685	Human immunodeficiency virus type 1 (ARV2/SF2 ISOLATE)		6		25		19
	11678	Human immunodeficiency virus type 1 (BH10 ISOLATE)		5		18		13
	11684	Human immunodeficiency virus type 1 (BH8 ISOLATE)		1		1		
	11693	Human immunodeficiency virus type 1 (BRAIN ISOLATE)		2		2		
	11686	Human immunodeficiency virus type 1 (BRU ISOLATE)		3		6		3
X	11687	Human immunodeficiency virus type 1 (CDC-451 ISOLATE)				3		3
	11679	Human immunodeficiency virus type 1 (CLONE 12)		2		7		5
	11689	Human immunodeficiency virus type 1 (ELI ISOLATE)		8		12		4
	11706	Human immunodeficiency virus type 1 (HXB2 ISOLATE)		28		40		12
	11707	Human immunodeficiency virus type 1 (HXB3 ISOLATE)		4		4		
X	362651	Human immunodeficiency virus type 1 (isolate YU2)				1		1
X	11694	Human immunodeficiency virus type 1 (JH3 ISOLATE)				3		3
	11688	Human immunodeficiency virus type 1 (JRCSF ISOLATE)		33		35		2
X	36375	Human immunodeficiency virus type 1 (KB-1 isolate)				1		1
	11697	Human immunodeficiency virus type 1 (MAL ISOLATE)		10		17		7
	11704	Human immunodeficiency virus type 1 (MFA ISOLATE)		2		2		
	11696	Human immunodeficiency virus type 1 (MN ISOLATE)		19		21		2
	11695	Human immunodeficiency virus type 1 (NDK ISOLATE)		3		10		7
	11698	Human immunodeficiency virus type 1 (NEW YORK-5 ISOLATE)		1		1		
	11699	Human immunodeficiency virus type 1 (OYI ISOLATE)				3		3
	11700	Human immunodeficiency virus type 1 (PV22 ISOLATE)		3		4		1
	11701	Human immunodeficiency virus type 1 (RF/HAT ISOLATE)		24		26		2
X	11691	Human immunodeficiency virus type 1 (SF162 ISOLATE)				1		1
X	11690	Human immunodeficiency virus type 1 (SF33 ISOLATE)				2		2

New 2006	NCBI TAX ID	SPECIES/STRAIN	B-05	T-05	B-06	T-06	Δ B	Δ T
	11703	Human immunodeficiency virus type 1 (STRAIN UGANDAN / ISOLATE U455)		1		14		13
	31678	Human immunodeficiency virus type 1 (WMJ1 isolate)		15		15		
	11705	Human immunodeficiency virus type 1 (WMJ2 ISOLATE)		1		2		1
	11683	Human immunodeficiency virus type 1 (Z2/CDC-Z34 ISOLATE)		1		3		2
	11708	Human immunodeficiency virus type 1 (ZAIRE 6 ISOLATE)		2		2		
	82834	Human immunodeficiency virus type 1 lw12.3 isolate		16		16		
X	11714	Human immunodeficiency virus type 2 (ISOLATE BEN)				7		7
X	11715	Human immunodeficiency virus type 2 (ISOLATE CAM2)				1		1
X	11713	Human immunodeficiency virus type 2 (ISOLATE D194)				1		1
X	11716	Human immunodeficiency virus type 2 (ISOLATE D205,7)				1		1
	11717	Human immunodeficiency virus type 2 (ISOLATE GHANA-1)		1		6		5
X	73484	Human immunodeficiency virus type 2 (isolate KR)				2		2
	11719	Human immunodeficiency virus type 2 (ISOLATE NIH-Z)		1		1		
X	11718	Human immunodeficiency virus type 2 (ISOLATE SBLISY)				2		2
	11721	Human immunodeficiency virus type 2 (ISOLATE ST)		1		1		
	10566	Human papillomavirus		2		2		
	10580	Human papillomavirus type 11		3		5		2
	333760	Human papillomavirus type 16	1	27	1	58		31
	333761	Human papillomavirus type 18		4		12		8
X	37112	Human papillomavirus type 29				1		1
	10585	Human papillomavirus type 31				7		7
X	10586	Human papillomavirus type 33				3		3
X	10592	Human papillomavirus type 44				1		1
X	10593	Human papillomavirus type 45				2		2
	10618	Human papillomavirus type 52		1		4		3
X	333765	Human papillomavirus type 53				1		1
X	10596	Human papillomavirus type 56				4		4
X	10598	Human papillomavirus type 58				3		3
	37115	Human papillomavirus type 59		1		1		
	10600	Human papillomavirus type 6b		5		5		
	10602	Human papillomavirus type me180		1		1		
X	36412	Human parainfluenza 1 virus (strains A1426 / 86-315 / 62M-753)				1		1
X	10798	Human parvovirus B19			90	16	90	16
X	12082	Human poliovirus 1 strain Sabin			1		1	

New 2006	NCBI TAX ID	SPECIES/STRAIN	B-05	T-05	B-06	T-06	Δ B	Δ T
	11260	Human respiratory syncytial virus A strain Long		3		3		
	11259	Human respiratory syncytial virus A2		1		2		1
X	10962	Human rotavirus (SEROTYPE 1 / STRAIN WA)			1		1	
	10960	Human rotavirus (SEROTYPE 4 / STRAIN ST. THOMAS 3)				1		1
	11927	Human T-cell lymphotropic virus type 1 (Caribbean isolate)		10		10		
	11926	Human T-cell lymphotropic virus type 1 (strain ATK)		2		3		1
	11908	Human T-lymphotropic virus 1	1	6	1	6		
X	9580	Hylobates lar				5		5
X	77644	IncQ plasmid pIE1120				1		1
	11320	Influenza A virus	7	37	8	130	1	93
X	11322	Influenza A virus (A/Alaska/6/77(H3N2))				1		1
X	150147	Influenza A virus (A/Aichi/2/68(H3N2))			7	9	7	9
X	11323	Influenza A virus (A/Anas acuta/Primorje/695/76(H2N3))				10		10
	135322	Influenza A virus (A/Ann Arbor/6/60(H2N2))		2		180		178
X	62446	Influenza A virus (A/Argentina/3779/94(H3N2))				1		1
	11325	Influenza A virus (A/Bangkok/1/79(H3N2))		11	1	4	1	-7
X	11327	Influenza A virus (A/Beijing/11/56 (H1N1))				22		22
X	62450	Influenza A virus (A/Beijing/32/92(H3N2))				54		54
X	150159	Influenza A virus (A/Budgerigar/Hokkaido/1/77 (H4N6))				4		4
	36418	Influenza A virus (A/chicken/Brescia/1902(H7N7))		1		1		
X	11384	Influenza A virus (A/chicken/FPV/Weybridge(H7N7))				1		1
X	97348	Influenza A virus (A/Chicken/Hong Kong/G23/97(H9N2))				1		1
X	11342	Influenza A virus (A/chicken/Pennsylvania/1370/83(H5N2))			1		1	
X	241242	Influenza A virus (A/Chile/1/83/(H1N1))				1		1
X	62541	Influenza A virus (A/Christ_Church/2/88(H3N2))				1		1
X	107493	Influenza A virus (A/Cordoba/3278/96(H3N2))				1		1
X	88295	Influenza A virus (A/duck/Alberta/35/76(H1N1))				2		2
X	11354	Influenza A virus (A/duck/England/1/56(H11N6))				1		1
	11363	Influenza A virus (A/duck/Hokkaido/10/85(H3))		1				-1
	258040	Influenza A virus (A/duck/Hokkaido/107/01(H2N3))		1				-1
X	80266	Influenza A virus (A/Duck/Hokkaido/8/80 (H3N8))				6		6

New 2006	NCBI TAX ID	SPECIES/STRAIN	B-05	T-05	B-06	T-06	Δ B	Δ T
X	210671	Influenza A virus (A/duck/Hong Kong/366/78(H9N2))				1		1
X	353253	Influenza A virus (A/duck/Novosibirsk/56/2005(H5N1))				2		2
X	11374	Influenza A virus (A/duck/Ukraine/1/63(H3N8))			3		3	
X	137578	Influenza A virus (A/England/939/69 x A/PR/8/34)				1		1
	135658	Influenza A virus (A/Environment/Hong Kong/437-4/99 (H5N1))						
X	229411	Influenza A virus (A/Fort Monmouth/1/47-MA(H1N1))				6		6
X	260806	Influenza A virus (A/FPV/Dutch/27(H7N7))			1		1	
X	191583	Influenza A virus (A/FPV/Rostock/34(H7N1))			2	4	2	4
X	107558	Influenza A virus (A/France/75/97(H3N2))				1		1
X	31661	Influenza A virus (A/Harbin/1/88(H1N2))				4		4
X	144556	Influenza A virus (A/Hong Kong(H3N2))				15		15
	108859	Influenza A virus (A/Hong Kong/1/68(H3N2))		2		2		
X	130760	Influenza A virus (A/Hong Kong/1073/99(H9N2))			1		1	
X	130763	Influenza A virus (A/Hong Kong/156/97(H5N1))				2		2
X	317652	Influenza A virus (A/Hong Kong/2/68(H3N2))				2		2
X	88104	Influenza A Virus (A/Hong Kong/483/97(H5N1))			1		1	
	220500	Influenza A virus (A/Japan/305/57(H2N2))	1	18	1	26		8
X	11422	Influenza A virus (A/Kiev/59/79(H1N1))				1		1
X	220985	Influenza A virus (A/Korea/426/68(H2N2))				1		1
	11428	Influenza A virus (A/Leningrad/54/1(H1N1))		1		3		2
X	62559	Influenza A virus (A/Los_Angeles/(H3N2))			5		5	
	107471	Influenza A virus (A/Louisiana/1/93(H3N2))						
	95888	Influenza A virus (A/mallard duck/PA/10218/84(H5N2))	1		5		4	
X	220503	Influenza A virus (A/Memphis/1/71(H3N2))			6	11	6	11
X	79695	Influenza A virus (A/Memphis/1/71H-A/Bellamy/42N)				1		1
	252233	Influenza A virus (A/Memphis/102/72(H3N2))		3		7		4
	228928	Influenza A virus (A/Memphis/31/98(H3N2))	4		6		2	
X	11440	Influenza A virus (A/Memphis/6/86(H3N2))			2		2	
X	62488	Influenza A virus (A/Nanchang/58/93(H3N2))				1		1
	132734	Influenza A virus (A/Netherlands/018/94 (H3N2))						
X	132841	Influenza A virus (A/Netherlands/785e/90 (H3N2))				1		1
	132842	Influenza A virus (A/Netherlands/889/91 (H3N2))		1		1		
X	335313	Influenza A virus (A/New York/364/2004(H3N2))				1		1
X	62564	Influenza A virus (A/New_York/15/94(H3N2))				1		1

New 2006	NCBI TAX ID	SPECIES/STRAIN	B-05	T-05	B-06	T-06	Δ B	Δ T
X	62496	Influenza A virus (A/New_York/17/94(H3N2))				1		1
	260805	Influenza A virus (A/NT/60/68/(H3N2))		14		64		50
X	155917	Influenza A virus (A/NWS/33HA-A/tern/Australia/G70C/75NA)			1		1	
X	170501	Influenza A virus (A/NWS-G70c(H1N9))			1		1	
X	62503	Influenza A virus (A/Ohio/3/95(H3N2))				1		1
	223935	Influenza A virus (A/Okuda/57(H2N2))		9		9		
	119209	Influenza A virus (A/Philippines/2/82(H3N2))	1		1			
	83922	Influenza A virus (A/Port Chalmers/1/73(H3N2))		1	7	1	7	
	107478	Influenza A virus (A/Puerto Rico/1/97(H3N2))	1				-1	
	211044	Influenza A virus (A/Puerto Rico/8/34(H1N1))	3	171	15	316	12	145
	183764	Influenza A virus (A/Puerto Rico/8/34/Mount Sinai(H1N1))	1	8	7	62	6	54
X	106423	Influenza A virus (A/Quail/Hong Kong/AF157/92(H9N2))				1		1
X	221012	Influenza A virus (A/Rio/6/69(H3N2))				1		1
X	260814	Influenza A virus (A/seal/Mass/1/80(H7N7))			1		1	
X	62512	Influenza A virus (A/Shangdong/5/94(H3N2))				1		1
X	63105	Influenza A virus (A/Shanghai/16/89(H3N2))				1		1
X	220949	Influenza A virus (A/Singapore/1/57(H2N2))			1		1	
X	150154	Influenza A virus (A/swine/29/37 (H1N1))				2		2
	169169	Influenza A virus (A/swine/Cotes d'Armor/1482/99(H1N1))		1		1		
X	11498	Influenza A virus (A/swine/Hong Kong/126/82(H3N2))				1		1
	219637	Influenza A virus (A/swine/Hong Kong/6/76(H3N2))		3		1		-2
	219641	Influenza A virus (A/swine/Hong Kong/81/78(H3N2))		2		2		
	145307	Influenza A virus (A/swine/Hong Kong/9/98(H9N2))	2		2			
	11504	Influenza A virus (A/swine/Indiana/1726/88(H1N1))		1		1		
	260809	Influenza A virus (A/swine/Iowa/15/30(H1N1))		2				-2
X	300743	Influenza A virus (A/swine/Korea/S10/2004(H1N1))				1		1
X	300744	Influenza A virus (A/swine/Korea/S175/2004(H1N1))				1		1
X	186460	Influenza A virus (A/swine/New Jersey/11/76(H1N1))			1		1	
	11517	Influenza A virus (A/swine/Ukkel/1/84(H3N2))		1		1		
	88305	Influenza A virus (A/swine/Wisconsin/1/61(H1N1))		2		1		-1
X	82372	Influenza A virus (A/Sydney/05/97-like(H3N2))				1		1

New 2006	NCBI TAX ID	SPECIES/STRAIN	B-05	T-05	B-06	T-06	Δ B	Δ T
X	11475	Influenza A virus (A/tern/Australia/G70C/75(H11N9))			10		10	
	183796	Influenza A virus (A/Texas/1/77(H3N2))	1	8	3	12	2	4
X	11470	Influenza A virus (A/turkey/Ontario/7732/66(H5N9))			5	6	5	6
	11480	Influenza A virus (A/Udorn/307/72(H3N2))		3		4		1
X	62596	Influenza A virus (A/USSR/26/(H3N2))				1		1
	225083	Influenza A virus (A/USSR/90/77(H1N1))	7	3	11	1	4	-2
	62522	Influenza A virus (A/Victoria/1/93(H3N2))		1				-1
	78586	Influenza A virus (A/Victoria/3/75(H3N2))		5	41	6	41	1
X	284217	Influenza A virus (A/Viet Nam/1194/2004(H5N1))				2		2
X	11484	Influenza A virus (A/whale/Maine/1/84(H13N9))			3		3	
X	11487	Influenza A virus (A/Wilson-Smith/33(H1N1))				108		108
	145855	Influenza A virus (A/WSN/33(H1N1))	2		6		4	
	63106	Influenza A virus (A/Wuhan/359/95(H3N2))	3		5		2	
	132504	Influenza A virus (A/X-31(H3N2))	2	4	46	143	44	139
	11357	Influenza A virus (STRAIN A/DUCK/HOKKAIDO/5/77)		1		1		
X	11408	Influenza A virus (STRAIN A/EQUINE/NEW MARKET/76)				1		1
	41857	Influenza A virus H3N2		4	1	16	1	12
	11520	Influenza B virus	1		2	4	1	4
	184816	Influenza B virus (B/Kadoma/122/99)	2		2			
X	256080	Influenza B virus (B/Kobe/1/2003)			1		1	
	107412	Influenza B virus (B/Lee/40)		2	5	3	5	1
X	150127	Influenza B virus (B/Osaka/983/97-V3)			1		1	
	107417	Influenza B virus (B/Singapore/222/79)		1				-1
	107418	Influenza B virus (B/Victoria/2/87)			2		2	
X	11521	Influenza B virus (STRAIN B/ANN ARBOR/1/86)				3		3
X	11531	Influenza B virus (STRAIN B/HONG KONG/8/73)				1		1
X	11541	Influenza B virus (STRAIN B/OREGON/5/80)			17		17	
	11553	Influenza C virus (C/Ann Arbor/1/50)	1		9		8	
	197911	Influenzavirus A		12		5		-7
X	9725	Inia geoffrensis				1		1
	42097	Isla Vista virus	1		1			
	11072	Japanese encephalitis virus	2		9	1	7	1
	11075	Japanese encephalitis virus strain JAOARS982	4	2	4	4		2
	11076	Japanese encephalitis virus strain Nakayama	2		2	1		1
	51240	Juglans regia	1		1			
	11619	Junin virus				1		1
	13101	Juniperus ashei	5		5			
	42894	Khabarovsk virus	1		1			

New 2006	NCBI TAX ID	SPECIES/STRAIN	B-05	T-05	B-06	T-06	Δ B	Δ T
	10638	Kilham polyomavirus						
	571	Klebsiella oxytoca						
	573	Klebsiella pneumoniae	15		15			
	11078	Kunjin virus (STRAIN MRM61C)		3		3		
X	11577	La Crosse virus				1		1
X	1360	Lactococcus lactis subsp. lactis				1		1
	33728	Lake Victoria marburg virus - Popp	1		1			
	33727	Lake Victoria marburgvirus - Musoke	1		4		3	
X	11620	Lassa virus			4	4	4	4
	11621	Lassa virus GA391		1		10		9
	11622	Lassa virus Josiah		4		66		62
	5664	Leishmania major		1		1		
X	1642	Listeria innocua			2	1	2	1
	1639	Listeria monocytogenes		20	17	49	17	29
	169963	Listeria monocytogenes EGD-e	5	34	5	41		7
X	265669	Listeria monocytogenes str. 4b F2365				3		3
X	217686	Little cherry virus 1				1		1
	4522	Lolium perenne	7	10	7	13		3
X	4081	Lycopersicon esculentum				1		1
	11623	Lymphocytic choriomeningitis virus		17	1	53	1	36
	11624	Lymphocytic choriomeningitis virus (strain Armstrong)		29	2	70	2	41
X	11625	Lymphocytic choriomeningitis virus (strain Pasteur)				1		1
	11627	Lymphocytic choriomeningitis virus (strain WE)		5		11		6
X	9541	Macaca fascicularis				3		3
X	10373	Macaca mulatta cytomegalovirus				35		35
	11628	Machupo virus						
	7130	Manduca sexta		2		2		
	45201	Mannheimia haemolytica serotype 1	13		13			
X	11234	Measles virus				10		10
	11237	Measles virus strain IP-3-CA		1		1		
	11239	Measles virus strain Yamagata-1		3		3		
X	12107	Mengo virus			1		1	
X	10036	Mesocricetus auratus			24		24	
X	11801	Moloney murine leukemia virus				1		1
X	300180	Mopeia Lassa reassortant 29				1		1
	11629	Mopeia virus		1		2		1
	11759	Mouse mammary tumor virus (STRAIN C3H)						
	11171	Mumps virus (STRAIN MIYAHARA VACCINE)						
X	10366	Murid herpesvirus 1			2	1	2	1
	33708	Murid herpesvirus 4		3		4		1
	10367	Murine cytomegalovirus (strain Smith)		2		2		
X	11138	Murine hepatitis virus				1		1

New 2006	NCBI TAX ID	SPECIES/STRAIN	B-05	T-05	B-06	T-06	Δ B	Δ T
	76344	Murine hepatitis virus strain 2						
	12760	Murine hepatitis virus strain 4		2		2		
X	11142	Murine hepatitis virus strain A59			1		1	
	11144	Murine hepatitis virus strain JHM		1		4		3
	11079	Murray Valley encephalitis virus		3	12	6	12	3
X	301478	Murray valley encephalitis virus (strain MVE-1-51)			10	10	10	10
	10090	Mus musculus	4	80	63	192	59	112
	10092	Mus musculus domesticus		12		12		
X	1763	Mycobacterium			2	1	2	1
	1764	Mycobacterium avium		2	2	2	2	
	1765	Mycobacterium bovis		21	17	70	17	49
X	233413	Mycobacterium bovis AF2122/97				30		30
	33892	Mycobacterium bovis BCG	7	37	7	91		54
X	1766	Mycobacterium fortuitum				1		1
	1767	Mycobacterium intracellulare	1		1	1		1
X	1768	Mycobacterium kansasii			7		7	
	1769	Mycobacterium leprae	22	51	36	181	14	130
	272631	Mycobacterium leprae TN		3		3		
	1771	Mycobacterium phlei		1		1		
	1773	Mycobacterium tuberculosis	13	279	115	555	102	276
	83331	Mycobacterium tuberculosis CDC1551		13		13		
	83332	Mycobacterium tuberculosis H37Rv	29	146	29	259		113
X	2097	Mycoplasma genitalium				1		1
X	35670	Naja naja			2		2	
	485	Neisseria gonorrhoeae		9		9		
	487	Neisseria meningitidis		1		1		
X	4097	Nicotiana tabacum			1		1	
X	121791	Nipah virus			4		4	
	122928	Norovirus genogroup 1	1		1			
	122929	Norovirus genogroup 2	1		1			
X	95340	Norwalk-like virus			1		1	
	42764	Oliveros virus				1		1
X	9733	Orcinus orca				1		1
	784	Orientia tsutsugamushi	3		89		86	
	9986	Oryctolagus cuniculus		3		5		2
	9940	Ovis aries		1	51	4	51	3
	1406	Paenibacillus polymyxa		1		1		
	9598	Pan troglodytes		13	4	15	4	2
X	747	Pasteurella multocida				8		8
	5076	Penicillium chrysogenum	10		10			
	6978	Periplaneta americana	3		3			
	15957	Phleum pratense	21		21			
X	9742	Phocoena phocoena				1		1
	9755	Physeter catodon		25		26		1

New 2006	NCBI TAX ID	SPECIES/STRAIN	B-05	T-05	B-06	T-06	Δ B	Δ T
	11630	Pichinde virus		4		4		
	3888	Pisum sativum		1		1		
X	141833	Plasmid pIPO2T				2		2
	5821	Plasmodium berghei		2	3	7	3	5
X	5825	Plasmodium chabaudi			2	2	2	2
	5833	Plasmodium falciparum	1	75	108	186	107	111
X	57266	Plasmodium falciparum (isolate 7G8)			1		1	
	5835	Plasmodium falciparum (isolate CAMP / Malaysia)		1	1	3	1	2
X	5837	Plasmodium falciparum (isolate FC27 / Papua New Guinea)			4	3	4	3
X	5839	Plasmodium falciparum (isolate K1 / Thailand)			5	6	5	6
X	70153	Plasmodium falciparum (isolate mad20 / papua new guinea)				11		11
	5843	Plasmodium falciparum (isolate NF54)		6	2	14	2	8
X	5842	Plasmodium falciparum (isolate NF7 / Ghana)			2		2	
	57270	Plasmodium falciparum (isolate PALO ALTO / UGANDA)		5	2	5	2	
X	5834	Plasmodium falciparum (isolate RO-33 / Ghana)			2	1	2	1
X	5846	Plasmodium falciparum (isolate T4 / Thailand)			1	1	1	1
	5848	Plasmodium falciparum (isolate WELLCOME)		1	2	15	2	14
	36329	Plasmodium falciparum 3D7	4	134	20	163	16	29
X	5857	Plasmodium fragile				4		4
X	5850	Plasmodium knowlesi			2	6	2	6
X	5851	Plasmodium knowlesi strain H			4	2	4	2
X	5852	Plasmodium knowlesi strain Nuri				3		3
	5854	Plasmodium reichenowi		1		6		5
	5855	Plasmodium vivax	3	4	7	4	4	
X	31273	Plasmodium vivax (strain Belem)			1		1	
	5861	Plasmodium yoelii		1		10		9
	73239	Plasmodium yoelii yoelii		4	10	13	10	9
X	352914	Plasmodium yoelii yoelii str. 17XNL			2		2	
X	5145	Podospora anserina				1		1
	138953	Poliovirus	1		1			
X	12088	Poliovirus type 3 (strains P3/LEON/37 AND P3/LEON 12A[1]B)			1		1	
X	9600	Pongo pygmaeus				2		2
	28344	Porcine respiratory and reproductive syndrome virus	1		1			
	11151	Porcine transmissible gastroenteritis coronavirus (STRAIN PURDUE)	1		1			
	837	Porphyromonas gingivalis		1		1		
	11603	Prospect Hill virus	1		1			

New 2006	NCBI TAX ID	SPECIES/STRAIN	B-05	T-05	B-06	T-06	Δ B	Δ T
X	584	<i>Proteus mirabilis</i>			1		1	
	287	<i>Pseudomonas aeruginosa</i>	3		23	34	20	34
X	294	<i>Pseudomonas fluorescens</i>				3		3
X	303	<i>Pseudomonas putida</i>				5		5
X	306	<i>Pseudomonas</i> sp.				1		1
X	74138	<i>Pseudomonas</i> sp. DJ-12				1		1
X	71238	<i>Pseudomonas</i> sp. G-179				1		1
X	159091	<i>Pseudomonas</i> sp. KIE171				1		1
X	237609	<i>Pseudomonas</i> sp. KL28				1		1
X	150396	<i>Pseudomonas</i> sp. MT-1				1		1
	11604	Puumala virus	1	1	38	2	37	1
X	11605	Puumala virus (STRAIN HALLNAS B1)				1		1
	39000	Puumala virus (strain k27)		6		6		
	39002	Puumala virus (strain sotkamo/v-2969/81)	131		196		65	
X	39003	Puumala virus (strain udmurtia/894cg/91)				1		1
X	11976	Rabbit hemorrhagic disease virus			1		1	
	11292	Rabies virus	2	2	12	10	10	8
X	11293	Rabies virus (strain AVO1)				3		3
X	11294	Rabies virus (strain CVS-11)			13		13	
	11295	Rabies virus (strain ERA)	4	6	9	18	5	12
	11296	Rabies virus (strain HEP-FLURY)	1	1	5	1	4	
	37132	Rabies virus (strain Ontario fox)	3		3			
	39005	Rabies virus (strain ontario skunk)		2		2		
	103929	Rabies virus (strain Pasteur / PV)	1	2	1	1		-1
	45418	Rabies virus (strain vnukovo-32)		1		1		
	10116	<i>Rattus norvegicus</i>	6	54	9	70	3	16
	12814	Respiratory syncytial virus		7		7		
X	103930	Rhesus cytomegalovirus strain 68-1				5		5
	160061	Ricinus		1		1		
	3988	Ricinus communis	3	3	8	3	5	
X	781	<i>Rickettsia conorii</i>				6		6
	272944	<i>Rickettsia conorii</i> str. Malish 7		5		5		
X	782	<i>Rickettsia prowazekii</i>				10		10
X	783	<i>Rickettsia rickettsii</i>				1		1
X	35793	<i>Rickettsia sibirica</i>				3		3
X	785	<i>Rickettsia typhi</i>				1		1
X	11588	Rift Valley fever virus			4		4	
X	11589	Rift valley fever virus (STRAIN ZH-548 M12)				1		1
	11241	Rinderpest virus	3		3			
	46920	Rio Mamore virus	1		1			
	37207	Rio Segundo virus	1		1			
X	11032	Ross river virus (STRAIN T48)			3		3	
X	10912	Rotavirus				1		1
	11041	Rubella virus	1		1			
	11045	Rubella virus (STRAIN THERIEN)		1		1		

New 2006	NCBI TAX ID	SPECIES/STRAIN	B-05	T-05	B-06	T-06	Δ B	Δ T
	45709	Sabia virus		1		11		10
	4932	Saccharomyces cerevisiae		3	1	2	1	-1
X	9521	Saimiri sciureus				1		1
	10381	Saimiriine herpesvirus 2				1		1
X	590	Salmonella			3		3	
X	29477	Salmonella enterica subsp. enterica serovar Essen			1		1	
X	70803	Salmonella enterica subsp. enterica serovar Minnesota			1		1	
X	599	Salmonella sp.			1		1	
	601	Salmonella typhi	2	6	6	18	4	12
	602	Salmonella typhimurium		2	1	34	1	32
	227859	SARS coronavirus	46	136	114	179	68	43
	228407	SARS coronavirus BJ01	23	19	25	19	2	
	228415	SARS coronavirus CUHK-W1	17		17			
X	229992	SARS coronavirus Frankfurt 1			1		1	
	227984	SARS coronavirus Tor2	51	5	51	1957		1952
	228330	SARS coronavirus Urbani	2		30	26	28	26
	6183	Schistosoma mansoni		1		1		
	4896	Schizosaccharomyces pombe		1		1		
	11033	Semliki forest virus	1		5		4	
	11191	Sendai virus		2		2		
	11194	Sendai virus (strain Enders)		1		9		8
	11198	Sendai virus (Z)		1		1		
X	12557	Seoul virus 80-39				7		7
	11610	Seoul virus SR-11	1		1			
X	615	Serratia marcescens				1		1
	4182	Sesamum indicum	11		11			
X	622	Shigella dysenteriae			2		2	
	623	Shigella flexneri	3	1	82	4	79	3
X	42897	Shigella flexneri 2a			26		26	
X	624	Shigella sonnei				1		1
	11723	Simian immunodeficiency virus		119		438		319
X	11711	Simian immunodeficiency virus - mac				1		1
X	31682	Simian immunodeficiency virus - mac1A11				8		8
	31683	Simian immunodeficiency virus - stm		3		3		
X	11737	Simian immunodeficiency virus (F236/SMH4 ISOLATE) (SOOTY MANGABEY)				7		7
	11735	Simian immunodeficiency virus (K6W ISOLATE)		50		72		22
X	11736	Simian immunodeficiency virus (K78 ISOLATE)				7		7
	11733	Simian immunodeficiency virus (MM142-83 ISOLATE)				50		50
X	11734	Simian immunodeficiency virus (MM251 ISOLATE)				1		1

New 2006	NCBI TAX ID	SPECIES/STRAIN	B-05	T-05	B-06	T-06	Δ B	Δ T
	11738	Simian immunodeficiency virus (PBJ/BC13 ISOLATE) (SOOTY MANGABEY)						
X	160753	Simian immunodeficiency virus 17E-Fr				2		2
	10633	Simian virus 40		4		7		3
	57667	Simian-Human immunodeficiency virus		1		1		
	37705	Sin Nombre virus	1		1	11		11
X	11780	Snyder-Theilen feline sarcoma virus			1		1	
	6584	Spisula solidissima		3		3		
X	7108	Spodoptera frugiperda				1		1
	1280	Staphylococcus aureus	17	4	79	25	62	21
X	93062	Staphylococcus aureus subsp. aureus COL			1		1	
X	359787	Staphylococcus aureus subsp. aureus JH1			1		1	
X	158878	Staphylococcus aureus subsp. aureus Mu50			1		1	
X	196620	Staphylococcus aureus subsp. aureus MW2				7		7
X	158879	Staphylococcus aureus subsp. aureus N315			14		14	
	1309	Streptococcus mutans	1		1			
X	1313	Streptococcus pneumoniae				2		2
	1314	Streptococcus pyogenes	4	8	57	79	53	71
X	160490	Streptococcus pyogenes M1 GAS			12		12	
X	286636	Streptococcus pyogenes MGAS10394				1		1
	301449	Streptococcus pyogenes serotype M5		1	50	62	50	61
X	301450	Streptococcus pyogenes serotype M6			5	4	5	4
X	36470	Streptococcus sp. 'group A'			13		13	
	1915	Streptomyces lincolnensis		1		1		
	31523	Suid herpesvirus 1 (strain Indiana-Funkhauser / Becker)		1		1		
	9823	Sus scrofa	1	4	1	4		
X	32630	synthetic construct				2		2
	11631	Tacaribe virus				1		1
X	13281	Taphozous georgianus				2		2
	12124	Theiler's encephalomyelitis virus		8		8		
	12127	Theiler's encephalomyelitis virus (STRAIN GDVII)						
X	271	Thermus aquaticus			1		1	
	11084	Tick-borne encephalitis virus	12		12	1		1
X	11087	Tick-borne encephalitis virus (STRAIN SOFJIN)			6		6	
	83192	Topografov virus	1		1			
X	7787	Torpedo californica			4		4	
	7788	Torpedo marmorata		3		3		
	5811	Toxoplasma gondii	7	8	20	24	13	16
	160	Treponema pallidum						
	9337	Trichosurus vulpecula		7		7		
	4565	Triticum aestivum		20		81		61
	5693	Trypanosoma cruzi	1	19	1	76		57
	37133	Tula virus	1		1			

New 2006	NCBI TAX ID	SPECIES/STRAIN	B-05	T-05	B-06	T-06	Δ B	Δ T
	32644	unidentified	4	79			-4	-79
	38018	unidentified bacteriophage		9		1		-8
	11309	unidentified influenza virus		12		14		2
	10245	Vaccinia virus		25		133		108
X	10251	Vaccinia virus (strain IHD-J)			1		1	
	126794	Vaccinia virus Ankara		2		10		8
X	10249	Vaccinia virus Copenhagen				100		100
	10254	Vaccinia virus WR		2	7	2869	7	2867
X	12870	Variola major virus				2		2
X	10255	Variola virus				1		1
X	11037	Venezuelan equine encephalitis virus (STRAIN TC-83)			21		21	
X	11038	Venezuelan equine encephalitis virus (STRAIN TRINIDAD DONKEY)			14		14	
X	7742	Vertebrata			24	1	24	1
	11276	Vesicular stomatitis virus		2		3		1
X	11278	Vesicular stomatitis virus (serotype Indiana / strain Glasgow)				1		1
	11284	Vesicular stomatitis virus (strain Orsay)						
	666	Vibrio cholerae	12		54	34	42	34
	127906	Vibrio cholerae O1	8		13		5	
	686	Vibrio cholerae O1 biovar eltor	3		3			
X	670	Vibrio parahaemolyticus				17		17
X	223926	Vibrio parahaemolyticus RIMD 2210633			2		2	
X	672	Vibrio vulnificus				26		26
	11288	Viral hemorrhagic septicemia virus (STRAIN 07-71)	1		1			
	3972	Viscum album	17		17			
	11082	West Nile virus	4	14	7	34	3	20
X	46919	Whitewater Arroyo virus				2		2
X	66077	Wolbachia sp. wMel				1		1
	10436	Woodchuck hepatitis virus w64 (ISOLATE PWS23)		1		1		
	8355	Xenopus laevis		7		13		6
X	8364	Xenopus tropicalis				1		1
X	11089	Yellow fever virus			1		1	
	11090	Yellow fever virus (STRAIN 17D)	1	9	1	9		
X	11091	Yellow fever virus (STRAIN PASTEUR 17D-204)			2		2	
	630	Yersinia enterocolitica		1	3	9	3	8
	632	Yersinia pestis	5	28	28	61	23	33
X	186538	Zaire ebolavirus				21		21
X	4577	Zea mays				1		1

2 Website Features

The IEDB website functionality can be divided into five categories – Query, Resources, Account Information, Support, and About IEDB, which largely correspond to the pull-down menus at the top of the home page. The subsections that follow describe the website features within these categories. These features can also be classified by those available to any user and those available only to registered users. Any user can become a registered user in a simple process described in Section 2.4.1. The features are listed in Table 2.

The reader will find it helpful to keep in mind how data are stored in the IEDB. Each item contained in the IEDB consists of a reference (article or submission) containing information about one or more epitopes and associated binding or response information. The same epitope can exist in multiple references. Therefore, the results of a query will contain the same epitope numerous times if the epitope is contained in multiple references.

Table 2 IEDB website features available to anonymous and registered users

Anonymous User Features	Registered User Features
<ul style="list-style-type: none"> • Tour the IEDB • Submit feedback • View online help • View home page • Perform an Advanced Query • Perform a Simple Query • Perform a Search • Download Advanced Query results • Download Simple Query or Search results • Utilize analysis tools • Utilize finders when performing queries • View forum discussion topics • Search forum discussion topics • Download export files • Download documents • View external links • Use all features of the Analysis Resource 	<p>All features for anonymous users PLUS:</p> <ul style="list-style-type: none"> • Submit a help request to the help desk • Save advanced queries to be used at a later time • Submit a discussion topic to a forum • Reply to a previously submitted discussion topic in a forum

2.1 Home Page

The IEDB Home Page is the default screen displayed when users enter the IEDB system. Besides providing a general description of the IEDB project, the home page displays system level status and notification of scheduled updates or maintenance. The page also contains a variety of Quick Links and other information including Summary Metrics and project related News and Updates.

Summary Metrics are displayed in the lower left side of the screen. These numbers are intended to be a gauge of the volume of data available in the system. The Summary Metrics provide the

number of References, Records, Distinct Structures, and Distinct Epitopes that have been curated. These terms are fully defined in Section 2.2.4.

2.2 Query

There are six ways users can find information in the IEDB - Search, Simple Query, Advanced Query, Browse by Species, Browse by Allele, and Browse by 3D Structure. Users can retrieve information from curated literature or electronic submissions. Electronic submissions will usually contain relevant unpublished and unpatented epitope data. These six methods are elaborated in the subsections below.

The main menu bar appears on all screens (Figure 1). On the right hand side of the menu bar is a search field. This *Search* will return records that contain the term(s) entered, as in a Google search.



Figure 1 Picture of main menu bar

The *Advanced Query* is based on a standard Query By Example (QBE) approach, which is a method of forming queries where a user can enter conditions for each data field they want included in the query. The *Advanced Query* allows users to define example criteria for each field in the system. As there are over 300 fields, the *Advanced Query* function is both powerful and comprehensive. *Advanced Query* results are displayed in columnar format similar to a spreadsheet. The system allows the user to select which columns are displayed in the results and to download data.

The *Simple Query* is also based on the QBE approach. This form however, displays only the fields necessary to perform popular queries. It is geared to help users get the quick answers to the questions they typically ask the IEDB and is easier to read and understand. The results of a *Search* or *Simple Query* are organized into various lists the user can pick from. This allows the user to drill down quickly to the level that is of interest.

2.2.1 Browse for Records

The IEDB allows users to browse for records in three different ways – by MHC allele, by source species, and by 3D structure. Browse by 3D structure is a new feature for the year.

2.2.1.1 Browse Records by MHC Allele

All users can find records associated with a specific MHC allele by browsing records by allele. To browse records by allele, the user selects *Browse Records by Allele* under the *Query* heading on the main menu or the Quick Links on the IEDB home page. A table will be displayed with all the MHC alleles contained in the IEDB, sorted alphabetically by allele name. The number of records associated with each allele will be shown as a link. Clicking on a link will display the records associated with the selected allele.

2.2.1.2 Browse Records by Source Species

Users can find records associated with a specific epitope source species by browsing records by species. To do this, the user accesses the *Browse Records by Species* page via the Query pull-down menu or the Quick Links on the IEDB home page. A table will be displayed with all the epitope source species contained in the IEDB. The number of records associated with each species will be shown as a link. Clicking on a link will display the records associated with the selected species.

2.2.1.3 Browse Records by 3D Structure

Users can find IEDB records associated with PDB structures by browsing records by 3D structure. To do this, the user accesses the *Browse Records by 3D Structure* page via the Query pull-down menu or the Quick Links on the IEDB home page. A table will be displayed that indicates the number of B Cell, MHC binding, and T Cell assays with structural data. The number of records/assays for each category will be shown as a link. Clicking on a link will display the associated records.

2.2.2 Perform a Search

A Search will allow all users to locate records in the database using a keyword, identifier, or sequence. Wild card characters '%' and '_' can be used in the search field. The '%' character will match zero or more characters and the '_' character will match exactly one character. Additionally, the operators 'and', 'or', and 'not' can be utilized. The search field is case insensitive, so "ABC" is the same as "abc". Regular expressions cannot be used in the search field.

To perform a search, users enter criteria such as a keyword, identifier, or sequence into the search field on the main menu (Figure 1). They then click the Search button and view the Result Summary, which is described in Section 2.2.4.

In addition to using the search function from the menu bar, users can submit searches via a properly constructed URL. The example below will perform a search using the URL for records that contain 'dengue'. This would return the equivalent of typing 'dengue' in the search option on the menu bar and clicking the Search button.

<http://www.immuneEpitope.org/httpQuery.do?dispatch=runquery&searchValue=dengue>

2.2.3 Perform a Simple Query

The simple query was designed to allow users to perform common queries without having to search through the numerous fields available on the Advanced Query. To access the Simple Query search form, users select *Perform a Simple Search* under the *Query* heading on the main menu. The system will display the Simple Query Search Form (Figure 2). Data in the IEDB consists of References, Structures, and Assays. A reference can have many structures and a structure can have many assays. Each reference will have at least one structure and each structure will have at least one assay. When users perform a search, the system will determine which records match their search criteria down to the assay since that is the lowest level in the IEDB data hierarchy. The search form is written with descriptive logic to help the user understand how providing search criteria will affect the results.

Return all assays that.....	
Have either:	<input checked="" type="radio"/> Positive Data Only OR <input type="radio"/> Any Data Present
AND the host/immunized species is:	<input type="text"/> <input type="button" value="Species Finder"/>
AND	<input checked="" type="checkbox"/> Involve T-Cell Data
	- T-Cell Details -
	Fall into one of the following assay categories:
	<input checked="" type="checkbox"/> MHC Binding <input checked="" type="checkbox"/> MHC Ligand Elution <input checked="" type="checkbox"/> T-Cell Response
	AND have the following MHC restriction:
	<input type="text"/> <input type="button" value="Allele Finder"/>
	OR
	<input checked="" type="checkbox"/> Involve B-Cell Data
	- B-Cell Details -
	Fall into the B-Cell Assay Category
	AND the antibody used is one of the following:

Figure 2 Top portion of the Simple Query Search Form

Some fields will allow multiple selections as search criteria. In these cases the selections are treated as a set. Records will be considered a match if they include at least one of the selected values in the set. Some fields will use finders (Section 2.2.6) to help users when the number of possible choices is extensive. For example the Species Finder and Allele finder are visible in the Simple Query Search Form seen in Figure 2.

If the T Cell Data section is selected, at least one of the assay categories (MHC Binding, MHC Ligand Elution, and T Cell Response) must be selected. If the B Cell Data section is selected, at least one of the antibody types (Monoclonal and Polyclonal) and at least one of the reference types (Literature, Submission) must be selected.

A Regular expression (<http://www.digitalmars.com/ctg/regular.html>) can be entered in the Linear Sequence field. Depending on the way a regular expression is written, users can search for single or multiple sequences.

The 'Citation or abstract contains' field at the bottom of the form utilizes the following wildcard characters: '%' and '_'. A '%' character matches zero or more characters and a '_' matches exactly one character. Also the operators 'And', 'Or', and 'Not' can be used with these fields. The following illustrates the precedence of the above operators from highest precedence to lowest: Not, And, Or. Parentheses can be used to alter precedence. Common words such as 'the', 'a',

'but', etc. are ignored if used as search criteria. Operators may be nested, but if no operators are used between words, results will include records that contain the entered string.

To perform a simple query, users select *Perform a Simple Query* under the *Query* heading on the main menu or in the Quick Links box on the IEDB home page. Users enter their search criteria, click the *Submit* button, and view the Results Summary (Section 2.2.4).

2.2.4 Search and Simple Query – Results Summary Page

The Result Summary page displays the results of a Search or Simple Query. The number of references, records, distinct epitopes, distinct structures, and assays (displayed as links) that met the specified search criteria are listed. The user can click on the number in the Count column (Figure 3) to view all the results for that category. The Result Summary allows the user to quickly assess the results of their search and drill down the level of detail that is of interest.



Record Type		Count
References		47
Records		331
Distinct Epitopes 		
	Peptide	238
	DNA, RNA	0
	Lipid, Carbohydrate, Glycolipid, Other	0
Distinct Structures 		
	Peptide	325
	DNA, RNA	0
	Lipid, Carbohydrate, Glycolipid, Other	0
Reported Measurements		
	MHC Binding Assays	229
	MHC Ligand Elution Assays	0
	T Cell Assays	456
	B Cell Assays	347

Figure 3 Sample of Result Summary Table

The number of **references** is the total number of references containing the matching records. Each **record** contains data about one epitope (or structure) in one reference. Any number of assays can be associated with a record. A protein/DNA sequence/carbohydrate etc. that may or may not induce an immune response or MHC binding is considered a structure. A structure is an

epitope if it produces a positive qualitative measurement for at least one of its associated assays. A structure is not an epitope if it produces a negative qualitative measurement for each of its associated assays. All epitopes are structures, but not all structures are epitopes. Multiple epitopes/structures may be described in a single reference. If the same epitope is described in two references, and a user defines criteria that match the epitope, the two records would be returned which are identical except for the reference information and possibly the assay information.

A **distinct epitope** is an epitope with given characteristics regardless of the reference. If the same epitope is described in two different references, and a user defines criteria that match the epitope, only one distinct epitope will be displayed in the results, although all associated assays from both references would be reflected in the reported number of assays. The number of distinct epitopes will always be less than or equal to the number of distinct structures.

A **distinct structure** is a structure with given characteristics regardless of the reference. If the same structure is described in two different references, and a user defines criteria that matches the structure, only one distinct structure will be displayed in the results. The number of records will always be greater than or equal to the number of distinct structures.

The distinct epitopes and distinct structures are categorized according to chemical type (e.g., protein/peptide, carbohydrate, and lipid).

Assays are divided by assay categories and are associated with records. Therefore, if the same epitope and associated assay are described in two references, and a user enters search criteria for the epitope, two identical assays (with the exception of the reference information) will be listed.

When users click on the number of references on the Result Summary screen, the system will display the **Reference List**. This list will display all the references that matched the user's search criteria. The columns in the list on the Reference List screen are not comprehensive. Enough data is listed to identify each reference uniquely, including authors, authors' affiliations, article title, year, PubMed ID, and journal name. To view all the information related to a given reference, click the corresponding Details link. The list can be sorted by clicking on the column headers. This will sort the list based on the values in that column. Clicking on the same column again reverses the sort order. PubMed identifiers are displayed as a hyperlink in the list. Clicking on the PubMed identifier in the system will open the PubMed citation in a new window.

When users click on the number of records on the Result Summary screen the system will display the **Record List**. This list will display all the structures that matched the user's search criteria. The columns in the list on the Record List screen are not comprehensive, but provide enough data on the reference and epitope (IEDB ID, name, chemical type, structure, and source) to identify each structure uniquely. To view all the information related to a given structure the user can click the corresponding Details link. Columns can be sorted as described above.

A user can click on the number of distinct structures or distinct epitopes on the Result Summary screen the system to display the **Distinct Structure List** or **Distinct Epitope List**, which list all the distinct structures/epitopes, respectively, that matched the user's search criteria. The columns

displayed include the number of corresponding records, the number of positive and negative responses (epitopes have zero negative responses by definition), chemical type, chemical structure information, and source species. As in the other lists, only a subset of the data is displayed in order to identify each structure or epitope uniquely, and the user can view all the information by clicking on the hyperlink in the "# Records" column.

The **Assay List** is displayed when users click on the number of assays on the Result Summary screen. This list will display all the assays that matched the user's search criteria. Only data relevant to the assay type are included. Because of subset of all available is included, the user can view all the information related to a given assay by clicking the corresponding Details link.

From the Assay List display, users can perform additional analysis using tools in the Analysis Resource. Users can analyze one or more rows on the same page of the results by checking the corresponding check boxes, selecting the type of analysis in the Analyze Selected Records box (Figure 4) and clicking the *Analyze* button. Users can select from T Cell epitope prediction, B Cell epitope prediction, population coverage, conservancy analysis, and homology mapping. Each tool offers independent online help or instructions.

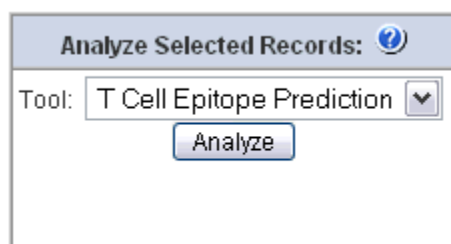


Figure 4 Analyze Selected Records box

The results of the reference, records, distinct epitope, distinct structure, and assay lists can be downloaded in a tab-delimited Excel format by clicking on the Excel link at the bottom of the screen. All rows will be downloaded regardless of the page shown or how many rows are displayed per page.

2.2.5 Advanced Query

The Advanced Search has numerous additional features that are not available through the Search and Simple Query methods. Most of these features are available to any anonymous user, but several are only available to registered users. Further information on these features is provided in the following subsections.

2.2.5.1 View Query Criteria

The criteria of a query will include the entered search criteria for the query. All users will have the ability to view the criteria of an Advanced Query after the query has been performed or after search criteria has been input into the Perform an Advanced Query form. Additionally, registered users will be able to view the criteria of their saved advanced queries.

2.2.5.2 Save Queries

Registered users can save a query after they have performed an Advanced Query. Users cannot

save a query after performing a Search or a Simple Query. A saved query will contain all of the entered filter criteria for the query and the format selected to report the results. When users save a query, they are required to give it a search title so it can be found in the future. Users might want to save a query so they can perform the query again at a later date when additional epitopes from other references have been added to the database. Once a query has been saved, users can load a saved query and then perform the query, or delete a saved query. Saved queries can also be modified, run, and saved again. Queries can be saved indefinitely if they are renewed by the user every year.

2.2.5.3 View Advanced Query Results

All users can view the results of an Advanced Query. To view the results of an Advanced Query, users must first perform the query for which they wish to view the results. The View Results page is displayed after an Advanced Query is performed which includes a Results Table. The Results Table will include epitope records with a limited amount of information and contains the following five columns: Links, Reference, Structure, Source, and Assay. One row will be displayed for each assay that meets the specified search criteria. Aggregate values for each assay will populate each column of the Results Table, except for the Links column, when the Summary format is selected. Under the Links column, each row in the table will have a Details link and possibly an Epitope Viewer link. Section 2.2.5.5 will discuss the Epitope Viewer link in further detail. Users can click on a Details link to view more detailed result information for the selected row. When a user clicks on a Details link, all populated reference fields, all populated epitope fields, and all populated binding or response fields for the corresponding assay will be displayed.

Most rows in the results display will typically not have an Epitope Viewer link. The user can map and visualize linear epitopes to homologous protein 3D structures from the Protein Data Bank by selecting the assay of interest, selecting the "Homology Mapping" tool in the "Analyze Selected Records" box (Figure 4), and performing the analysis on the pages that follow.

Users can format the Results Table to include columns other than the default columns. Users have the option to select from several predetermined formats using the [change format](#) function, or to create a custom format. By default, the Summary format is used. Users also have the option to alter the number of rows displayed per page in the Results Table.

2.2.5.4 Download Advanced Query Results

All users can download the results of an Advanced Query in a tab-delimited format that can be read by Excel. To download the results of a query, users must first perform the query for which they wish to download the results. The columns that are downloaded are the columns displayed in the Results Table when the user selects the Excel link.

2.2.5.5 Accessing the EpitopeViewer

All users are able to utilize the EpitopeViewer after an Advanced Query has been performed. The EpitopeViewer is an application for three dimensional viewing of receptor-antigen interactions that can be accessed from the View Results screen. The EpitopeViewer can be used with all assays that have receptor-antigen interaction data available.

2.2.6 Finders Overview

Several finders (Allele, Assay, Disease, Source, and Species) are available to help facilitate selections and control vocabulary usage (improves result outputs). At times the potential list of selections can be quite extensive, and the finders help users make selections from large lists. All finders can be utilized when performing an Advanced Query. The allele, source, and species finders can also be used when performing a Simple Query. Multiple selections can be made when utilizing finders during a query. Records will be returned that contain one of the entered values, or at least one of the entered values in the case of the MHC Types Present field.

The disease and species finders use wild card characters by default on both ends of entered search criteria when the Search names button is selected. An exact match is found when criteria are entered and the Search ID's button is selected. The allele finder uses wild card characters by default on both ends of criteria entered into the allele field. The source finder uses wild card characters by default on both ends of criteria entered into the name or strain fields. All finder search fields can search on common strings such as 'the', 'and', 'a', 'or', 'but', etc.

2.2.6.1 MHC Allele Finder

The MHC Allele Finder facilitates the selection of one or more MHC alleles. Initially the Allele Finder lists all alleles ordered by allele name. The Allele Finder allows the user to find alleles by name and or class {I, II, non-classical} in the Find box (Figure 5). After the user supplies their search criteria and clicks the Search button, the system will filter the list of MHC alleles using the name and or class provided. The system then returns any alleles that contain the value in the name field and match the class selected.



The image shows a search form titled "Find:". It contains two input fields: "Class:" with a dropdown arrow and "Allele:" with a text box. Below the input fields are two buttons: "Search" and "Reset".

Figure 5 Search form on the MHC Allele Finder

2.2.6.2 Assay Type Finder

The Assay Type Finder is used to facilitate the selection of one or more assay types and lists all assay types in the selected assay category. The Assay Type Finder allows the user to find assay types by assay type name, assay group (formerly response measured) and or units in the Find box (Figure 6). After the user supplies their search criteria and clicks the Search button, the system filters the list of assay types using the selections provided in the Find box.

Figure 6 Search form on the Assay Type Finder

2.2.6.3 Disease Finder

The disease finder is used to facilitate the selection of a disease state and input the selection into a Disease State field. It includes all diseases from The International Statistical Classification of Diseases and Related Health Problems, tenth revision (ICD-10) and displays diseases with their corresponding ICD-10 codes in a hierarchical tree. The first level of the tree displays similar groups of diseases, and each additional level of the tree further breaks down the groups of diseases. Variations of each disease are not included. Searching capabilities are provided so users can quickly select a disease state.

The Disease Finder will allow the user to find diseases using the disease name or ICD-10 code (Figure 7). When the user performs a search, the system will display the first match and then allow the user to move forward and backward through the matching records using Next and Previous buttons (Figure 8). When the user provides a name, any disease name that contains the character string provided will be considered a match.

Figure 7 Search form on the Disease Finder

The selections in the disease finder are displayed in a tree (Figure 8). To expand a node of the tree, the user clicks the plus sign next to the name. To collapse a node, a user clicks the corresponding minus sign. The ICD-10 code is displayed next to each selection in square brackets. For example the ICD-10 code for Bartonellosis in the Figure 8 is A44.

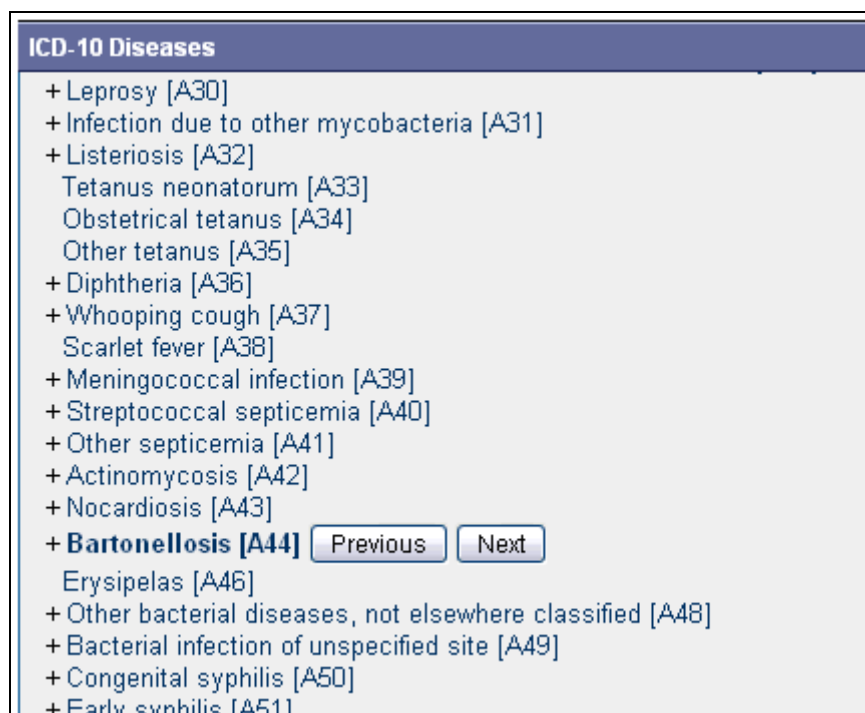


Figure 8 Disease List

2.2.6.4 Source Finder

The Source Finder is used to facilitate the selection of source antigens, immunogens, and epitopes. Records in the Source Finder come from GenBank, UniProt and IEDB curators. Among the finders, the Source Finder has the most comprehensive and flexible search form, and even includes the Species Finder. Due to the large volume of possible selections, the Source Finder initially won't display any selections. Users need to perform a search to narrow the list down. After the user enters search criteria (Figure 9) and clicks the Search button, the system will list the matching sources from which the user may select, as seen in Figure 10. The user can then select their desired sources from the list by click on *Select* in the far left column.

Find:

Accession/GI No.: Database:

Chemical Type: Antigen Nature:

Name:

Source Species:

Strain:

Figure 9 Search form on the Source Finder

3,681 items found, displaying 1 to 10.

Pages [First/Prev] [1](#), [2](#), [3](#), [4](#), [5](#), [6](#), [7](#), [8](#) [Next/Last]

Options	Accession	Database	Name(s)	Source Species	Antigen Nature	Chemical Type
Select	38479432	GenBank	surface layer protein SapA8 ; sapA8	Campylobacter fetus subsp. fetus		Peptide/P
Select	83031685	GenBank	PB1-F2 protein; PB1-F2	Influenza A virus (A/Puerto Rico/8/34(H1N1))		Peptide/P
Select	10048255	GenBank	malaria exported protein-1; EXP-1	Plasmodium falciparum		Peptide/P
Select	10048261	GenBank	sporozoite surface protein 2; SSP2	Plasmodium falciparum		Peptide/P
Select	1017427	GenBank	elastic titin; titin	Homo sapiens		Peptide/P
Select	1042034	GenBank	1-phosphatidylinositol-4-phosphate 5-kinase isoform C; PtdIns4P 5-kinase isoform C; 1-phosphatidylinositol-4-phosphate 5-kinase isoform C; PtdIns4P	Homo sapiens		Peptide/P

Figure 10 Source List

2.2.6.5 Species Finder

The species finder is used to facilitate the selection of a species or virus from the NCBI Taxonomy Database. Common selections are displayed initially to speed the selection of the usual suspects, as seen in Figure 11. However, if the species of interest is not on the “common selections” default list, the user can search the NCBI Taxonomy tree by clicking the NCBI Taxonomy Tree in the accordion slider. The Common Selection bar will move down to reveal the taxonomy tree.

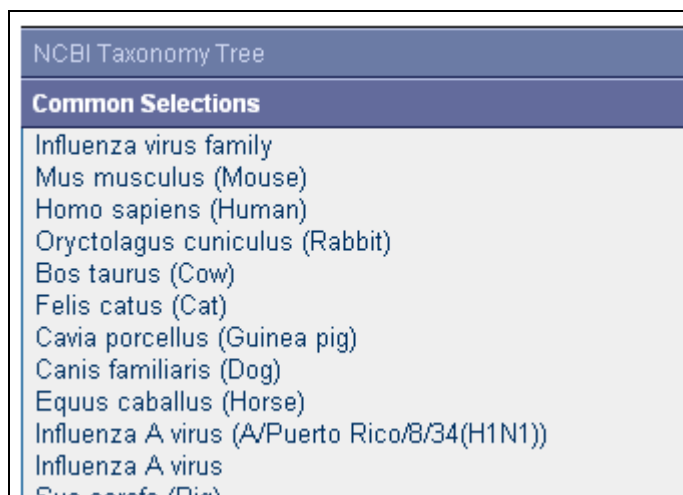


Figure 11 Common selections in Species Finder

The Species Finder will allow the user to find species by their name or taxonomy identifier (assigned by NCBI). The user performs the search by typing the name or taxonomy identifier in

the search form at the top of the page, similar to the one seen in Figure 7 for the Disease Finder. The system will display the first match and then allow the user to move forward and backward through the matching records using Next and Previous buttons (Figure 12). The system will consider a match achieved when the user-supplied character string for the species name or synonym is contained in an entry in the taxonomy tree.

The selections in the species finder are displayed in a tree (Figure 12). The taxonomy identifier is displayed next to each node of the tree in square brackets. For example, the taxonomy identifier for the selection in the example below is 301536. To see the synonyms for a selection, users can place their computer mouse over the scientific name.

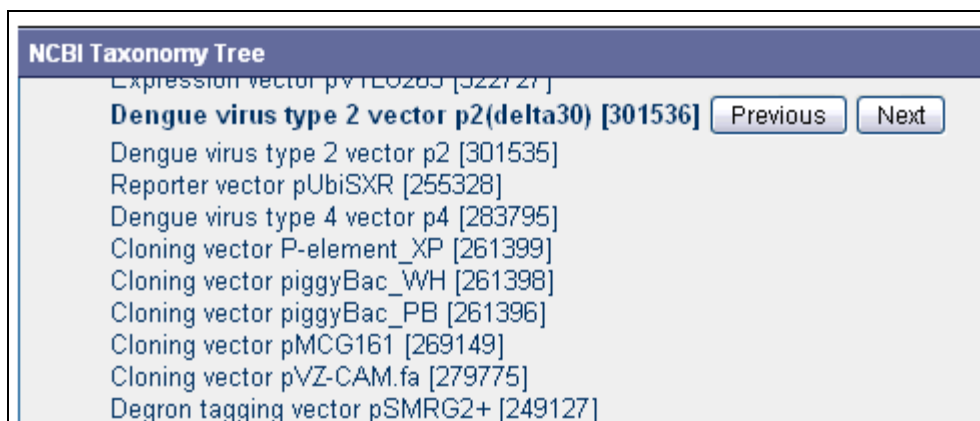


Figure 12 NCBI Taxonomy Tree

2.2.7 Analysis Tools

All users will be able to utilize analysis tools after an Advanced Query, Simple Query, or Search has been executed. From the View Results screen (for advanced queries) or an Assay List page (for simple queries and searches), the user will have the ability to perform various types of analysis on results. The analysis tools are described in greater detail in Section 2.3.1.1

2.3 Resources

There are five major types of resource features available to users. Users can utilize the analysis and epitope prediction capabilities of the Analysis Resource. They can also make use of the IEDB forum to discuss and exchange information and ideas. In addition, users can export the contents of the database for their own use and manipulation, and download an ever increasing number of documents from the website. Finally, links to other websites of interest and relevance are provided.

2.3.1 Analysis Resource

The purpose of the Analysis Resource of the IEDB is to provide computational tools that enhance the value of the IEDB database to the user. Providing access to tools in one centralized location helps make users aware of available solutions to their problems. All of the information contained within the Analysis Resource, including analysis tools and algorithms developed by the IEDB staff, are freely available to the scientific community.

The tools provided in the analysis resource fall into two categories – analysis tools and predictive tools. Analytical tools help extract and interpret data contained in the database. Currently four tools exist - population coverage (Section 2.3.1.1.1), conservancy analysis (Section 2.3.1.1.2), epitope visualization (Section 2.3.1.1.3), and mapping B-cell epitopes to the PDB structures (Section 2.3.1.1.4).

Predictive tools extrapolate beyond data held in the database. They can be used to predict epitopes in protein sequences or predict properties of known epitopes, such as their MHC binding affinity. One of the fundamental challenges of predictive tools is determining the accuracy and validity of their outputs. An evaluation of three predictive tools for MHC class I binding was conducted in 2006 (Peters et al., PLoS Comput. Biol., 2006) to help address this issue.

For predictive tools, it is important to differentiate between the **tool** making predictions, and the **method** used to generate that tool, given a set of training data. For example, the artificial neural network method or approach, when trained on a particular data set, will yield a predictive tool. As the ANN method is trained on different data sets, different corresponding predictive tools will result. In this way, the ANN method can be used to develop an MHC class I prediction tool and a separate MHC class II prediction tool. These tools can be refined as more data are available for training. One benefit of the IEDB is that it allows implementing methods to automatically generate new predictive tools as the database grows.

Predictive tools can be subdivided into categories by what they aim to predict. The current tools fall into the subcategories listed below:

- Proteasomal cleavage
- TAP transport
- Binding to MHC class I / II
- B-cell epitope prediction

The next subsections will describe each analytical and predictive tool in more detail.

2.3.1.1 Analysis Tools

2.3.1.1.1 Population coverage

T cells recognize a complex between a specific MHC type and a particular pathogen-derived epitope and thus a given epitope will elicit a response only in individuals that express an MHC molecule capable of binding that particular epitope. MHC molecules are extremely polymorphic (over a thousand different variants are known in humans). Therefore, selecting multiple peptides with different MHC binding specificities will afford increased coverage of the patient population targeted as vaccine recipients. The issue of population coverage in relation to MHC polymorphism is further complicated by the fact that different MHC types are expressed at dramatically different frequencies in different ethnicities. Thus, without careful consideration, a vaccine with ethnically biased population coverage could result. To address this issue, the actual/predicted binding capacity of potential epitopes to as many different MHC molecules

possible (and when available, also restriction data of T cell responses recognizing the epitope) can be used to project the population coverage in different ethnicities of different vaccine candidates or epitope sets. Accordingly, epitope-based vaccines or diagnostics can be designed to maximize population coverage, while minimizing complexity (that is, the number of different epitopes included in the diagnostic or vaccine), and also minimizing the variability of coverage obtained or projected in different ethnic groups.

An important consideration in the process of epitope selection is that the patient population coverage afforded by a given set is not simply corresponding to the sum of the coverage of its individual components. Thus, to calculate the coverage afforded by a given mixture of epitopes, a more comprehensive approach and a suitable algorithm has been developed for this specific purpose (Bui et al. BMC Bioinformatics 2006). This method calculates the fraction of individuals predicted to respond to a given epitope set on the basis of HLA genotypic frequencies, assuming non-linkage disequilibrium between HLA loci, and on the basis of MHC binding and/or T cell restriction data. The algorithm is briefly explained here. First, genotypic frequencies of various MHC are tabulated. Each time a peptide binds to a given MHC, a “hit” is recorded for that MHC. The process is repeated for all peptides. Then the hits for MHC are tallied. Next, the frequency of each possible diploid MHC combination (phenotype) is calculated. For n MHC types, this corresponds to an $n \times n$ tabulation of the frequency at which each specific pair of MHCs will be found in the population from which the MHC frequencies are derived. A similar table is generated to contain the number of hits per each of the MHC combinations by adding the number of hits associated with each of the two alleles of MHC in the combination (a simple exception is the case of homozygous combinations, where the number of hits is simply the number of hits of the given MHC). From these two tables, a frequency distribution is assembled, tabulating the genotypic frequency of all MHC combinations associated with a certain number of hits. The result of the analysis is displayed as a frequency distribution histogram and a cumulative frequency plot.

We have derived HLA allele genotypic frequencies from the dbMHC database (<http://www.ncbi.nlm.nih.gov/mhc/>) and stored them in a database on the IEDB tool server. At present, dbMHC provides allele frequencies for 78 populations and 11 different geographical areas. It is envisioned that the compiled data will be updated regularly as further HLA frequency data become available. Furthermore, customized frequency data can be utilized in the calculation, should studies of specific and particular patient populations be of interest to a given user. Multiple population coverages can be simultaneously calculated and an average population coverage is generated. Since MHC class I and II restricted epitopes elicit immune responses from two different T cell populations (CTL and Th cells, respectively), the program provides three different coverage calculation modes – (1) class I separate, (2) class II separate, and (3) class I and class II combined.

2.3.1.1.2 Epitope conservancy

In a diagnostic or epitope-based vaccine setting, focusing on conserved epitopes allows for targeting responses around pathogen variability, whether it exists prior to infection, or develops in the natural course of disease. The use of conserved epitopes would be expected to focus the immune response on sequences crucial for retaining biological function of the pathogen proteins, and thus with intrinsically lower variability, even under immune pressure. The epitope

conservancy analysis tools implemented here aims to address the issue of variability (or conservation) of epitopes, and to assist in the selection of epitopes with the desired pattern of conservation. The algorithm has been implemented to calculate the degree of conservancy of an epitope within a given protein sequence set at different degree of identities. The degree of conservation is defined as the number of protein sequences that contain the epitope at a given identity level, divided by the total number of protein sequences found in the dataset analyzed.

2.3.1.1.3 Visualization Tool

The convenient and easy to use EpitopeViewer, a Java application running JOGL, has been developed for three-dimensional visualization of immune epitopes and analyses of their interactions with antigen-specific receptors of the immune system (antibodies, T cell receptors, MHC molecules) for structures available in the Protein Data Bank (PDB). The EpitopeViewer is based on the Molecular Biology Toolkit (MBT; <http://mbt.sdsc.edu/>) developed at the San Diego Supercomputer Center (SDSC). It uses data both from the PDB and the IEDB, and visualizes one epitope at a time from a particular PDB structure.

The EpitopeViewer provides the following functionality:

- Link to the PDB web-page displaying a particular structure.
- Visualization of the 3D structure of epitope/antigen in complex with immune receptor(s) as curated within the IEDB and available in the PDB.
- Visualization of the 3D structure of epitope and antigen mapped to a PDB structure using the EpitopeMappingToPDB tool.
- Visualization of sequences of epitope/antigen and immune receptor(s).
- 3D-visualization of intermolecular (epitope-paratope, epitope-antibody CDR, epitope-MHC, pMHC-TCR, pMHC-TCR CDR), inter-atom and inter-residue interactions curated within the IEDB and/or calculated on the fly from the PDB file with essential details (contact type, atoms, distance) provided;
- 2D-plot of inter-residue interactions between epitope and immune receptor.
- Generation of publication-quality pictures of structures, sequences, and plots of contacting residues.

2.3.1.1.4 Tool for mapping linear epitopes to the PDB structures

EpitopeMappingToPDB is an automated tool that maps linear epitopes to three-dimensional protein structures within the PDB by sequence similarity of antigen and proteins in the PDB. For input, the tool uses the SwissProt ID of the antigen protein, the epitope sequence, and the position of the epitope in the antigen sequence as curated within the IEDB or input by the user. The tool applies the NCBI BLAST algorithm for performing sequence homology search, and provides options for the sophisticated user to choose cutoff values on parameters used in the search programs (such as e-value and penalty on gap initiation and gap extension). The tool

output page displays the alignment between the query sequence of the antigen containing the epitope and the sequence from the PDB representing significant hits (matches). The region within the epitope is highlighted in the alignment, and the sequence identity for the epitope and homologous region is provided. The EpitopeViewer application for visualization of homologous epitope/antigen and its further structural analysis is launched from the output page.

2.3.1.2 Predictive tools

2.3.1.2.1 MHC binding prediction methods

The Analysis Resource provides tools for predicting peptide binding to MHC class I and II molecules. For class I binding predictions, users can select predictions performed with tools derived from three different methods – artificial neural network, average relative binding, and stabilized matrix. Only the average relative binding method has been implemented for class II binding predictions. Tutorials and example data are available for both the class I and II tools.

2.3.1.2.1.1 Artificial Neural Network

Artificial neural networks (ANN) are computer algorithms modeled after the brain. They consist of many simple processing units which are wired together in a communication network. Each unit is a simplified model of a neuron which sends off a new signal if it receives a sufficiently strong input signal from the other units to which it is connected. The strength of these connections can be varied in order for the network to perform a desired pattern of node signal activity, which is learned from a set of input training data. The training data in this case are peptide sequences with quantitative affinities for a specific MHC molecule.

Many different implementations of artificial neural networks exist. The one utilized here is described for HLA-A2 binding predictions by Nielsen et al. (Protein Science, 2003) and has been applied to a number of different alleles (<http://www.cbs.dtu.dk/services/NetMHC/>).

2.3.1.2.1.2 Average Relative Binding (ARB)

Average relative binding (ARB) matrix binding prediction method is based on the assumption that each residue along the peptide molecule independently contributes to binding affinity. When a residue R occurs at position i in the peptide, it is assumed to contribute a constant amount of R_i to the free energy of binding of the peptide. The effect of each of the 20 possible amino acids at each possible position along the peptide sequence, therefore, can be estimated by a matrix of coefficients. The overall binding propensity of each peptide sequence, an algorithm “score”, is calculated by multiplying the R_i coefficients. Predicted IC_{50} values, which provide quantitative K_D (IC_{50}) predictions, are then calculated by mathematical transformations of the algorithm scores (Bui et al., Immunogenetics 2005).

2.3.1.2.1.3 Stabilized Matrix Method (SMM)

The Stabilized Matrix Method (SMM) described by Peters and Sette (BMC Bioinformatics, 2005) can be applied to calculate matrices from quantitative affinity data of peptides binding to MHC molecules. The advantage of this method is that it suppresses the noise present in the training data, caused by the inevitable experimental error as well as the limited number of data points.

2.3.1.2.1.4 MHC Class I Binding Prediction Resource

One of the new features introduced in 2006 was the MHC Class I Binding Prediction Resource (<http://mhcbindingpredictions.immuneepitope.org>). This site contains training data, test data, and other resources for tool developers interested in predictions of peptide binding to MHC class I molecules. The user can follow links to a manuscript describing the resource in detail (Peters et al., PLoS Comput. Biol. 2006), a dataset of experimental affinities of peptide to MHC molecules, and a description of the framework used for the evaluation of prediction methods. A link to this site appears on the Analysis Tools page in the T Cell Epitope Prediction box.

As described in the manuscript, predictions were obtained from public web-servers for all relevant peptide-MHC affinities in the dataset. The correlation between predicted and measured affinities was evaluated using scatter plots, linear regression, and ROC analyses. The evaluation of these external tools can be accessed on the site by name of the method or the MHC allele. As carefully noted in the manuscript, this is not a fair evaluation of the value of each method, primarily because the data available to each method are highly divergent.

A similar evaluation of the prediction performance of three prediction methods available in the IEDB Analysis Resource (ANN, ARB, and SMM) was carried out using cross-validation on the dataset. In contrast to the comparison of external predictions, this is a fair evaluation of prediction performance of the three methods, as training and testing data were the same for each method. Again, the evaluations of these three internal methods can be accessed by name of the method or the MHC allele.

2.3.1.2.2 *Proteasomal cleavage and TAP transport predictions*

For the prediction of antigen processing through the MHC class I antigen presentation pathway, we incorporated predictions of proteasomal cleavage and TAP transport similar to the MHCPathway website described in [Tenzer et al, CMLS, 2005]. The predictions are based on in vitro experiments characterizing the sequence specificity of proteasomal cleavage and TAP transport. The goal of the prediction is to identify MHC-I ligands (peptides that are naturally processed from their source proteins and presented by MHC class I molecules).

The proteasomal cleavage predictions evaluate how efficiently a peptide or its N-terminally prolonged precursors can be liberated from its source protein. The TAP transport predictions evaluate how efficiently a peptide or its N-terminal prolonged precursors are transported into the ER by TAP (Peters, J Immunol, 2003). When this information is taken together and combined with MHC class I binding predictions, the tool yields a prediction of the efficiency with which a peptide is presented on the cell surface.

2.3.1.2.3 *Sequence based antibody epitope prediction*

Five different tools are provided that predict antibody epitope candidates from amino acid sequences. Parameters such as hydrophilicity, flexibility, accessibility, and antigenic propensity of polypeptides chains have been correlated with the location of continuous epitopes in a few well-characterized proteins. Based on these observations, amino acid property scales have been developed to predict antigenic determinants. Each scale consists of 20 values assigned to each of the amino acid residues on the basis of their relative propensity to possess the property described by the scale. The following amino acid property scales have been selected and implemented based on their popularity and coverage of different categories.

- Hydrophobicity/hydrophilicity
 - Parker hydrophilicity prediction
- Flexibility
 - Karplus and Schulz flexibility prediction
- Surface exposure
 - Emini surface accessibility prediction
- Antigenicity
 - Kolaskar and Tongaonkar antigenicity prediction
- Secondary structure
 - Chou and Fasman beta turn prediction

Several new prediction tools for both continuous and discontinuous antibody epitopes are currently under development and will become available in the IEDB Analysis Resource during 2007.

2.3.2 Forums

The IEDB system includes a discussion forum to promote discussion, improve our user experience, and encourage community outreach. All users can search for discussion topics in forums, but only registered users can submit and reply to a discussion topic in a forum. Four forum categories were set up during the year – General, Queries/Reporting, Analysis Tools, and NIAID B Cell Epitope Prediction Tools Workshop. New forums will be created by an Application Administrator in response to feedback, help requests, and frequently asked questions. Apart from the IEDB system, the forum has its own help content provided by Jive Forums. The forums can be accessed at <http://www.immuneepitope.org/jive/index.jspa>.

2.3.3 Database Export

The contents of the Immune Epitope Database are exported on a regular basis to files in XML format that can be downloaded by any user. The Database Export screen lists one or more files that can be downloaded (Figure 13). To prevent the files from getting too large for download, the system can segment the entire download into multiple files as needed. Only the file or files for the latest database export is listed. Users can learn more about the IEDB XML format in the reference materials in the downloads area.

Export List	Size
IEDB_2006_5_2_0_30_23.xml	39056k
IEDB_2006_5_2_0_40_24.xml	5214k

Figure 13 Database export files

The file names are displayed as hyperlinks. The naming convention is IEDB_Year_Month_Day_Hours_Minutes_Counter.xml. The counter is needed to keep the files unique when more than one file is needed to represent a single export of the database. The counter does not repeat or reset for each export so the first file of each export won't start with

one. For example the first file above in Figure 13 was created May 2nd, 2006 at 12:30am. It was the 23rd export file created historically. Users can download the file by right-clicking on the corresponding hyperlink and selecting "save link as" or "save target as", depending on the browser.

2.3.4 Downloads

A variety of IEDB reference materials are available for download by all users. The Downloads page lists the files available for download by category. The files available for download include an Introduction to IEDB and Analysis Resource, the Curation Manual, IEDB Ontology, IEDB Annual Workshop Executive Summaries, the Annual Compendium, and quarterly newsletters. Additional reference materials will be added for download over time.

2.3.5 View External Links

The IEDB system provides a list of links to external resources solely for the convenience of Immune Epitope Database visitors. The Immune Epitope Database has no interest in, responsibility for, or control over the linked-site. The Immune Epitope Database makes no promises or warranties of any kind, express or implied, including those of fitness for any particular purpose, as to the content of the linked-site. To view the links available, select Links under the Resources heading on the main menu. The hyperlinks on the links page are grouped by category:

- Public Databases, Prediction Algorithms, and Other Tools
- Antibody Related Links
- Biodefense Resource Centers
- MHC and TCR Related Links
- Protein Related Links
- Laboratory Resources
- Biodefense Resources

2.4 Account Information

2.4.1 Become a Registered User

IEDB has two types of users: anonymous users and registered users. Anonymous users are persons who use the IEDB application and do not register with IEDB. Users can become registered with the IEDB by submitting an e-mail address and a password, and indicating the institution they belong to, the country where they live, and if they would like a newsletter subscription. Registered users are able to use features that are unavailable to anonymous users, such as submitting help requests or posting on discussion forums.

2.4.2 Registered User Login

Users must be registered with IEDB in order to log in to the system. Users will be automatically prompted to log in to the IEDB upon attempting to use a function that requires user registration.


2.4.3 Modify Account Information

Account modifications consist of e-mail address, institution, country, password, and/or news letter subscription changes. Users must already be registered in order to have an account to modify.

2.4.4 Terms of Use

The terms of use statement can be viewed by selecting *Terms of Use* under the *Account* heading on the main menu. The Terms of Use page is a collection of statements that outline the conditions related to using the IEDB system. The Terms of use includes our privacy notice, copyright information, and various disclaimers.

2.5 Support Overview

The IEDB offers numerous ways to learn how to use the IEDB and to receive help if a problem is encountered. On almost every page of the IEDB, instructions are listed in a box at the top of the screen that describes how to perform the most important functions on the page. The instructions can be removed from all pages by clicking on the Hide Instructions link in the Quick Links box on the home page. Additionally, a more detailed description of each page can be accessed by clicking the  button to the right of a page title. When performing an Advanced Query, a hint for each field can be seen by moving the cursor over a field name. The hint defines the field, describes how to enter a value, and gives an example. The user can click on a field name for a more detailed description of the field. The online help can also be accessed by moving the cursor over the Support tab and selecting Online Help, which describes all features of the IEDB in detail.

Numerous tours that describe how to perform different types of queries and provide an overview of the IEDB are available in the "About IEDB" pull-down menu. Additionally, the forums described in the Resources section (2.3.2) are available so users can view, post, and reply to discussion topics relating to the IEDB or immunology in general.

The IEDB has done its best to make the application intuitive and to describe how to use all features of the IEDB. However, questions may arise that the online help, tours, and forums do not answer. In these cases, users can submit feedback or submit a help request.

2.5.1 Submit Feedback

All users are able to submit feedback, which will transmit an email message to the IEDB team. Feedback is intended for questions, input, and suggestions, such as new features they would like to see added in the future. Feedback helps the IEDB team update the system to provide users with the best possible experience. If users need help using the system or handling an unexpected result, a Help Request is probably more appropriate.

2.5.2 Submit Help Request

Only registered users logged in to the system can submit help requests. Help requests will be sent to the help desk. Help requests should consist of problems that users have with the application, such as a certain function of the system not working, and not questions on how to use features of the application. The online help is designed to answer questions regarding

features of the application. A full name, day phone number, and description of the problem are required to submit a help request. After submitting a help request, a confirmation e-mail will be sent to the user's registered e-mail address, which will include the help request number. When the help request issue has been resolved, an e-mail will be sent to notify the user that the help request has been satisfied.

Help requests are generally responded to within one business day. Purely technical requests are often responded to and resolved in the same one day period. Requests that are specific to analytical tools or the method used to curate data are answered initially to inform the requestor that their question/comment is being forwarded to team specialists, and an approximate date of full response is provided. Based on the complexity of the request or if that request prompts the team to make changes to the system or curate additional data/source organisms, requests are resolved immediately in some cases, while others are resolved in future IEDB system builds, or later curation.

2.5.3 View Online Help

All users are able to view the online help. Online help is designed to be an online user guide. It describes in detail all of the features that the system offers, explains how users go about doing everything the system has to offer, and tries to answer any questions that users might have when using the system.

The Contents tab is selected by default. The Table of Contents is displayed on the left and a help topic is displayed in the section on the right. Users can click on a folder in the Table of Contents and it will expand. Help topics contained within the selected folder will appear. Users can click on a topic and the corresponding topic will appear in the section on the right. The contents pane organizes help topics as they might appear in a user manual.

Users can select the Index tab to view similar topics grouped together and displayed alphabetically. Users can also select the Search tab and enter criteria into the search field. Relevant topics will be displayed. The index view organizes help topics as they might appear in the index of a book.

2.6 About IEDB

The "About IEDB" pull-down menu was added after the initial beta release of the system. The "Tour the IEDB" menu item was moved from "Support" to this new location. Four other menu items were added – "Acknowledgements", "Publications", "Citing the IEDB", and "Powered By".

2.6.1 Tour the IEDB

Several video tours are available for viewing and download. They describe the IEDB in general, how to perform different types of queries, and how to interpret your query results. These tours are available in Flash Media format and are listed by subject. The title, length, and description for each video clip are listed. The script for each tour is provided in Microsoft Excel format. The tours and their descriptions are listed in Table 3.

Table 3 Flash media tours describing the basic IEDB functionality are available

Tour Category	Tour Name	Description
Overview	IEDB Overview	Provides an overview of the Immune Epitope Database and Analysis Resource or IEDB. The following modules are discussed: Query, Analysis tools, Discussion forum, and Online help.
Query Development Overviews	Search	Describes how to perform a Search.
	Simple Query	The Simple Query input form and how to input search criteria are discussed in this tour.
	Advanced Query	Outlines search criteria development and how registered users save advanced queries and load saved queries.
Query Results Overviews	Search and Simple Query	The result summary page and how to analyze and download results are discussed for a Search or Simple Query in this tour.
	Advanced Query	Describes how to format, download, and analyze Advanced Query results.
Sample Queries	Advanced Epitope Structure Query	Uses regular expression syntax to retrieve epitopes with a specific linear sequence using an Advanced Query. In this query, a regular expression is entered into the Epitope Linear Sequence field.
	Advanced Epitope Structure and Source Query 1	Demonstrates how to retrieve epitopes with a specific linear sequence and source name using an Advanced Query. The search fields include Epitope Linear Sequence and Epitope Source Name.
	Advanced Epitope Structure and Source Query 2	Demonstrates how to retrieve peptide epitopes for a species and use the custom report to format query results. The search fields include Epitope Chemical Type and Epitope Source Species.
	Advanced B Cell Query	Demonstrates how to find all epitopes with B Cell Response assays using monoclonal antibodies that react to a particular protein from a specific species strain using an Advanced Query. The search fields include B Cell Antibody Type and Epitope Source Name.
	Simple B Cell Query	Demonstrates how to find all epitopes with B Cell Response assays using monoclonal antibodies that react to a particular protein from a specific species strain using a Simple Query. The search fields include B Cell Antibody Type and Epitope Protein or Gene.
	Simple T Cell Query	Demonstrates how to retrieve T cell epitopes with a given MHC restriction using a Simple Query. The Assay Categories and MHC Restriction fields are used.

2.6.2 Acknowledgements

A host of talented individuals have worked hard to make the Immune Epitope Database a reality. A roster of the current team members can be viewed on this page.

2.6.3 Publications

A list of publications relevant to the IEDB can be found on the Publications page.

2.6.4 Citing the IEDB

Data and tools within the IEDB are presented as a public resource. Users are requested to consider citing the IEDB when they present information obtained from the IEDB or use tools contained in the Analysis Resource. It is expected that the authors of an entry as well as the IEDB are properly cited whenever their work is referred to:

1. The IEDB website should be cited using the URL: www.immuneepitope.org
2. The journal reference for the IEDB should be cited as:

Peters B, Sidney J, Bourne P, Bui HH, Buus S, Doh G, Fleri W, Kronenberg M, Kubo R, Lund O, Nemazee D, Ponomarenko JV, Sathiamurthy M, Schoenberger S, Stewart S, Surko P, Way S, Wilson S, Sette A. The immune epitope database and analysis resource: from vision to blueprint. *PLoS Biol.* 2005 Mar;3(3):e91

3 Scientific Publications

This section lists the scientific publications in 2006 for which the IEDB played a contributory role. Although citations of papers published by the IEDB team appeared during the year, neither a PubMed search nor a Google Scholar search provided any references outside the IEDB team that indicated that the IEDB played a contributory role. As a result, this year's list only includes IEDB scientific papers authored by the IEDB project team:

- Bui HH, Sidney J, Dinh K, Southwood S, Newman MJ, Sette A. Predicting population coverage of T-cell epitope-based diagnostics and vaccines. *BMC Bioinformatics*. 2006 Mar 17;7(1):153. PMID: 16545123
- Peters B, Bui HH, Frankild S, Nielson M, Lundegaard C, Kostem E, Basch D, Lamberth K, Harndahl M, Fleri W, Wilson SS, Sidney J, Lund O, Buus S, Sette A., A community resource benchmarking predictions of peptide binding to MHC-I molecules, *PLoS Comput. Biol.* 2006 Jun 9;2(6):e65. Epub 2006 Jun 9. PMID: 16789818
- Vita R, Vaughan K, Zarebski L, Salimi N, Fleri W, Grey H, Sathiamurthy M, Mokili J, Bui HH, Bourne PE, Ponomarenko J, de Castro R Jr, Chan RK, Sidney J, Wilson SS, Stewart S, Way S, Peters B, Sette A. Curation of complex, context-dependent immunological data. *BMC Bioinformatics*. 2006 Jul 12;7:341. PMID: 16836764
- Zhu S, Udaka K, Sidney J, Sette A, Aoki-Kinoshita KF, Mamitsuka H., Improving MHC binding peptide prediction by incorporating binding data of auxiliary MHC molecules, *Bioinformatics*. 2006 Jul 1;22(13):1648-55. Epub 2006 Apr 13. PMID: 16613909
- Salimi N, Vita R. The biocurator: connecting and enhancing scientific data. *PLoS Comput Biol.* 2006 Oct 27;2(10):e125. No abstract available. PMID: 17069454
- Sette A, Peters B. Immune epitope mapping in the post-genomic era: lessons for vaccine development. *Curr Opin Immunol.* 2006 Nov 16; PMID: 17113275
- Sette A, Bui, HH, Sidney J, Bourne P, Buus S, Fleri W, Kubo R, Lund O, Nemazee D, Ponomarenko JV, Sathiamurthy M, Stewart S, Way S, Wilson SS, Peters B, The Immune Epitope Database and Analysis Resource, *Pattern Recognition in Bioinformatics: International Workshop, PRIB 2006, Hong Kong, China, August 20, 2006 : Proceedings*, Springer, Jagath C. Rajapakse, Limsoon Wong, Raj Acharya (Eds). ISBN: 3540374469 9783540374466

4 References

- Bui HH, Sidney J, Peters B, Sathiamurthy M, Sinichi A, Purton KA, Mothe BR, Chisari FV, Watkins DI, Sette A. Automated generation and evaluation of specific MHC binding predictive tools: ARB matrix applications. *Immunogenetics*. 2005 Jun;57(5):304-14. Epub 2005 May 3.
- Bui HH, Sidney J, Dinh K, Southwood S, Newman MJ, Sette A. Predicting population coverage of T-cell epitope-based diagnostics and vaccines. *BMC Bioinformatics*. 2006 Mar 17;7(1):153. PMID: 16545123
- Nielsen M, Lundegaard C, Worning P, Lauemoller SL, Lamberth K, Buus S, Brunak S, Lund O., Reliable prediction of T-cell epitopes using neural networks with novel sequence representations, *Protein Sci*. 2003 May;12(5):1007-17
- Peters B, Bui HH, Frankild S, Nielson M, Lundegaard C, Kostem E, Basch D, Lamberth K, Harndahl M, Fleri W, Wilson SS, Sidney J, Lund O, Buus S, Sette A., A community resource benchmarking predictions of peptide binding to MHC-I molecules, *PLoS Comput. Biol*. 2006 Jun 9;2(6):e65. Epub 2006 Jun 9. PMID: 16789818
- Peters B, Sette A., "Generating quantitative models describing the sequence specificity of biological processes with the stabilized matrix method." *BMC Bioinformatics* 2005 May 31;6(1):132
- Peters B, Bulik S, Tampe R, Van Endert PM, Holzhtter HG. Identifying MHC class I epitopes by predicting the TAP transport efficiency of epitope precursors. *J Immunol*. 2003 Aug 15;171(4):1741-9.
- Tenzer S, Peters B, Bulik S, Schoor O, Lemmel C, Schatz MM, Kloetzel PM, Rammensee HG, Schild H, Holzhtter HG. Modeling the MHC class I pathway by combining predictions of proteasomal cleavage, TAP transport and MHC class I binding. *Cell Mol Life Sci*. 2005 May;62(9):1025-37.