### The Comprehensive Functional Annotation of Mouse Genes and Gene Products using The Gene Ontology (GO)



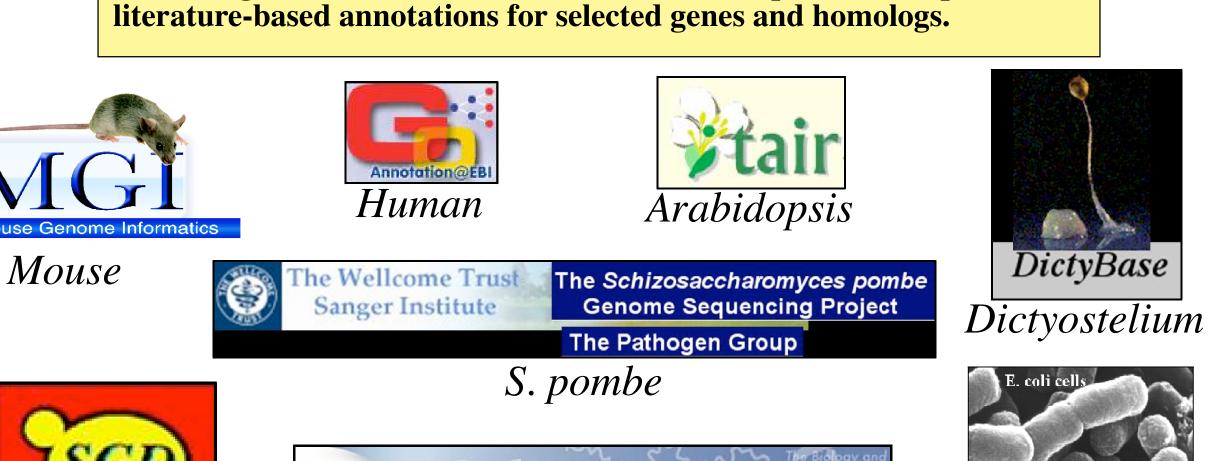


### Li Ni, Mary E. Dolan, Alex D. Diehl, Harold J. Drabkin, David P. Hill, Dmitry Sitnikov and Judith A. Blake The Jackson Laboratory, Bar Harbor, ME 04609

The Mouse Genome Informatics Database (MGI), integrates genetic, genomic, biological, and phenotypic data about the laboratory mouse to support the discovery of the genetic basis of functional mechanisms underlying heritable diseases. The Gene Ontology (GO), is a set of three structured vocabularies used by model organism databases such as MGI to identify the roles gene products play in the life of an organism. GO provides an ontology-driven functional annotation system that facilitates high-quality gene annotation for all species.

The GO Reference Genome Project is a shared annotation effort among the GO Consortium members who provide annotations for the nine primary model organisms such as mouse, fly and yeast. Starting with the set of genes implicated in human disease processes, the GO curators with the Model Organism Databases are coordinating their efforts in providing comprehensive annotations for the human disease genes and their orthologs. As the curators are simultaneously working on the same set of genes, they are also updating the ontologies, providing an orthology set for these organisms, and improving documentation of the GO annotation processes. The comprehensive annotations of the well-studied model organisms provide broad and deep annotation of the reference genomes and serve as a basis for the annotation of emerging genomes via sequence similarity matrices. Here we report progress for the MGI component of the GO Reference **Genome Project.** 

#### **Reference genomes**



Model organism databases coordinate efforts to provide comprehensive

#### **Target genes**

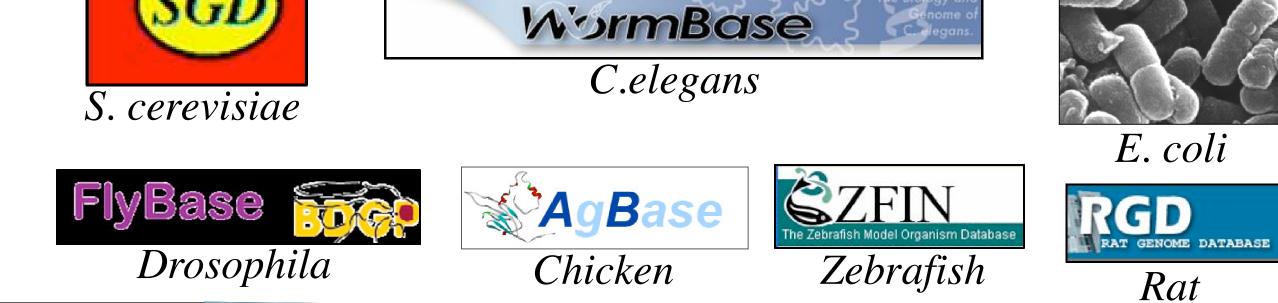
Each month a set of target genes is chosen for annotation. The criteria for selection are: 1. Implicated in human disease; 2. Part of a biochemical or signaling pathway; 3. Current "hot" genes; 4. Conserved homology across many species. As of October 2007, 273 genes have been selected. One selected gene, human *MSH6*, has homologs in nine other species. **Organism** <u>Gene</u> Msh6 Mus musculus **Msh6-predicted Rattus musculus Caenorhabditis elegans** msh-6 Dmel/CG7003 **Drosophila melanogaster** Dictostelium discoideum MSH6 Donio rorio mahh

The Gene Ontology project is supported by NHGRI grant HG0002273. The MGI project is supported by NIH grants HG000330.

#### **GO** annotations at MGI

The Mouse Genome Informatics (MGI) Database provides integrated access to data on the genetics, genomics and biology of the laboratory mouse. The gene detail page is the starting point to access data for a particular gene, including GO annotation.

_	(?)	Gene Detail	
IGI		Msh6	
NGLHOTO HOD	D	mut5 hamalag G (E. coll) NGI 13-3051	
		GTEF, GENED, Mehl Charansane 17	
e se jine		7.0 14 Detailed CenetroMaria 1, M	
more st. beleficines	-	Menning (Late) +)	
is on TC/ AppC/Homon Ciscaso Fup Wssion	Seipence Map		
of Wogle minel futtionary		(From NCBC an notation of NCBC Sulid 30) Ensembli Conticvice: LCSC Browson: ACSI Map Vice-or II II II III III III III III III III	
nepulo (ologui (HII)		MGENERAR (Pro-sa no an log bresi - (Manualar Orbitopi)	
ch Categories		Crimmenal ve Men (Muused) (Fish 10 (M)) Philen Super and Fisher directed on ten Msh5	
a na an ba - Chaige ang <b>MAC</b>	Марталан к	Pepirosensative Sequences	
/Marillars Lepes/Abelas -/Augustantsus		<ul> <li>☐ generik (17588) NEBI Gener Model MGI Becklened Deter (19980) C575L1X → U Ko</li> <li>☐ transition (1000000000000000000000000000000000000</li></ul>	
e-inn Da-s		Experipeptice PS (27.) UniFrot   EBC   MCC Sequence Detail (1958 Not Applicable)	
ara ian Majiwi ya si Mana/ Cata Mana/ Pintay		For the selected sequences due load integration at E. C.	
-/Climes miles	Phenotones	Al Nampania-(21) Al phonotypic ale cs( <u>a</u> ) : Targeted, mock cut( <u>a</u> ) Targeted, other <u>1</u> )	
dary terwaara Toti a tedro ara a tero biya (teti	Polymanithisms	RFLD(2) SNPp within Ska(2 from abSNP Suid 123)	
ar tosaalaa (CMIN) Titajia Cotoroy (Ni	, (60)		
ein à mediande l'ins 17 201	classifications	Commenten - Mind alma in gla modaan door lato Fondoor - Alle Lindoor stor Lindoo. All 60 desentet ens <u>(20)</u>	
(Druesa)	1 хроновии	GC (Farst ne in la (1) - of CAN understa(12)	
na wile	(III er delabase	EncompliGene Model <u>ENSKUS 300000005570</u>	
	c	Gene Ontology Classifications	
atore for mouse gen - Odi ment fil Still 2	niniary lest () et have assigned the 1960 – extern flars	ero Io=ing annotations to the gene product of MahCl (This text reflects annotations as ch Trusial jane is the illebra commeter right hould intrevial fleck the as of MasCl 2007. Terus	
anne mots homolog Historis ins Su aters for mouse ger - October 11, 2017 many contains ins anne fuel and We St 100510 ouggeste appin Miet e and th	rumany lest (1 let have assigned the MC - call of offers (match regarding the cal, Mad - Chetra Ma c shat Man) is hyper	rolo=ing annotations to the gene product of WahCl (This text reflects annotations as of	
anne mots homolog Historia (2011) annotations as Su ators for motse gen - Colored II (2012) anne fuel land We Si 100510 ouggeste appen Kiele and Historia other also	rimany lest (1 et have assigned the MC - restore films (match regarding the match regarding the cost, Mad. Chet all Ya cost, Mad. Chet all Ya cost at NSAC Sinvolve a Jan CC and Factors	erolo=ing annotations to the gene product of MbFC (The text reflects annotations at ch outse gene sciencialered commeter roboticul finite directions as of MeRC (2012). Terus simples gene <u>process of up kno=</u> . Fleate supply mouse gene cymbol and a PupMed ID. org thet all and flegen (the all reing the alleles KGEMC of FED 14+ KGEMC direction) and edit mismoton report General interaction with ACSI MSE, ACSI MSE, and ACSI MSE in reported	
anne mots homolog Historis as Su aters for mouse gen - October 11, 2017 many contonal mo anne fuel and We St 100510 ouggeste applin Miet et and th githe alle	rimary lest (1 et have assigned the 5 MC - restore films (match regarding the match regarding the cost, Mad. Cost all wat a that Nanbis involve a jan CC and success (notations in Tabu	Stolbeing annotations to the gene product of MbPC (The text reflects annotations at of upwarpenals, investment and of hour him exite the as of MeRC (2017) Takes since a product of MbPC (The text reflects annotations at of upwarpenals, investment and of hour him exite the as of MeRC (2017). The since a more taken by the set of the text reflects annotations at of the since and the period of the text reflects gene by the land a PubMed ID. On the text reflects of the text reflects MoD of text reflects and a PubMed ID. On the text reflects of the text reflects MoD of text reflects and a PubMed ID. On the text reflects of the text reflects MoD of text reflects and a PubMed ID. On the text reflects of the text reflects MoD of text reflects and text reflects of the and text reflects of the text reflects of text reflects of the text reflects of text reflects of the text reflects of the text reflects of the text reflects of the text reflects of text reflects of the text reflects of text refl	Delica
anne mots homolog Historia (2011) annotations as Su ators for motise gen - Colored III 2012 2013 any accretate (2012) anne fuel and We si 100510 auggeste anne fuel alor We colora, ne colora, ne col	rimary lest () et have assigned the have assigned the match relation () match relation () match relations () a lat Math Chetral Va c that Nanb Shrup ve a jan (Chetral Jacutes no tations in Tabu () Classification Term	erolowing annotations to the gene product of McHC (This text reflects annotations as of maximised in meillenear animeter of a function each all text merods of Merol 2012 of ani- single erologies of up know. Fleate supply mease gene cymbol and a PupMed ID. On the tail and Hagen Otian all using the abelas KoB MC of ECC 14 KoB MC direct of 11 and edin mismatch repair Generatinte rotion with AGSI MSC AGSI MSC and AGSI MSC in response is that Mehrie interface in tionand remark Research the fuel and walk and the Air Mark Mar Form (Text View) (CO Graph)	
anne mots homolog Historia (2011) attend for mouse gen - October (1000) anne foel can Wie anne foel can we anne foel can Wie anne foel can we foel can we foel can we	rimary lest (1 et have assigned the h MC - restore flow match regarding the match regarding the cost, Mach Chet all Ya- cost, Mach Chet all Ya- cost, Mach Chet all Ya- cost Nano Sinto Ya- algan OC all All-success motations in Tabu (1) Classification Term Classification (2) Classification	Enclosing an otet or site the gene product of MSEC (The test reflects an otet or site of investigate schedule and the dimensional feature as in Mast 2012, Theos structures gene present the sector structures gene combol and a PubMed ID. Enclose gene present the gene of a schedule feature with the MGEMC direct of the and the PubMed ID. Enclose gene present the gene of a schedule feature with the MGEMC direct of the and the PubMed ID. Enclose gene present the gene of a schedule for the Schedule feature with the MGEMC direct of the and the PubMed ID. Enclose gene present the schedule for the Schedule feature with the MGEMC direct of the and the Schedule feature for the Schedule feature from the Schedule feature for the Schedule feature from the Schedule feature feature feature from the Schedule feature feature from the Schedule feature feature feature from the Schedule feature	6733
anne mots homolog Historia (2011) annotations as Su accreter House gen - Coloner (110) 117 x any occretana (110) x any occretana (11	rimary lest (1 et have assigned the h MC - restore flow (match regarding the match regarding the cost, Mach Chetral Va- cost, Mach Chetral Va-	Probeing annotations to the gene product of MaRC (The test reflects annotations at of upwarpenals, includes, connects of high finite dial base mass of March 2017, fram, since segme process of up know Fleate suppriments and the time dial base measure with the NGLMC direct 21 fram, since segme process of up know Fleate suppriments and the Aleian March 2017, fram, since segme process of up know Fleate suppriments and the Aleian March 2017, fram, since segme process of up know Fleate suppriments and the Aleian March 2017, fram, since segme process of the Aleian March 2017, fram, since segme process of the Aleian March 2017, fram, since segme process of the Aleian March 2017, fram, since segme process of the Aleian March 2017, fram, since segme process of the Aleian March 2017, fram, since segme process of the Aleian March 2017, fram, since segme process of the Aleian March 2017, framework and the Aleian March 2017, framewor	5 1025 2 2:000
anne mots homolog Historis as Su attent for mouse gen - Coloren III 2002 Sub-state and We sub-state and We s	rimary lest (1 et have assigned the h MC - restore first match regarding the match regarding the match regarding the a cal, Mach Chet all va- c that Mach Chet all va- that that Mach Chet all va- that Mac	Problemg annotations to the gene product of MbPC (The text reflects annotations at cf         Invariants Considers commetation from the adia threatine as of Mas CONTROL Taro, is not set gene processe introduction in the adia threatine as of Mas CONTROL Taro, is not set gene processe introduction in the adiate MoN CONTROL TAR MON CONTROL TARO, is not the gene Docase interval on the adiate MoN CONTROL TARO MONORED AT the adiate threating the adiate monotonic set gene processe interval on the adiate MoN CONTROL TARO MONORED AT the adiate threating the adiate monotonic set gene processe interval on the adiate MoN CONTROL TARO MONORED AT the adiate threating the adiate monotonic set gene by the adiate threating the adiate monotonic set and the adiation of the adiate monotonic set adiate threating the adiate monotonic set gene by the adiate threating the adiate text of text o	5 16222 2:000 16111
anne mots homolog Historia (Serie) annotations as Su ators for motis gen - Colored III 2002 anne fuel ann We Si 100510 ouggeste anne fuel ann We Si 100510 ouggeste Si	rimary lest (1 techave assigned the h MC - restore flow (match regarding the match regarding the match regarding the match regarding the a chat Mark Chernal Va- a part Charateness motations in Tabu Classification for Al-Chernalism and STO process and STO process and STO process and STO process and STO process and Stormatic and Stor	<ul> <li>Prolowing annotations to the gene product of MarC (The text reflects annotations as of invariants, invalides, commeta and the finite set all text measy of MarC 2002 (Text), from since segme processional definition in the set of text measy of MarC 2002 (Text), from since segme processional definition in the set of text segme protect of the mark from text of text and tagen 00 and reining the shellar. Not NO 200200000 (Text) and NO 200200000000000000000000000000000000</li></ul>	10200 2:000 0:11 2:000
anne mots homolog Historia as Su abore for mouse ger - Coloren Hilbert 2012 2013 any accretate 100 anne fuel and We si 100010 augeste apon Kiele aler 100 anne fuel and We si 100010 augeste apon Kiele aler 100 anne fuel colore, ne colore, ne c	rimary lest () et have assigned the h MC - restore flax (match regarding the match regarding the match regarding the cost, Med. Cost allow cost, Med. Cost allow cost, Med. Cost allow cost Nando Sorteours a per OC at allow et hat Nando Sorteours () Cost allow cost allow () Cost allow () Cost allow () Cost allow () Cost allow () Cost allow () Cost allow () () Cost allow () () Cost allow () () () () () () () () () ()	<ul> <li>Polowing annotations to the gene product of MahC (The text reflects annotations as of inskelpenes, meillered to meter induct for each alterative as if MarC (1012) for a structure as inference on meter induction each alterative as if MarC (1012) for a structure as inference on the text of the structure as if MarC (1012) for a structure as inference on the structure and we as and we as an inference on the structure with VGSI MSC. NGSI MSC inference on the structure as infere</li></ul>	10200 2000 0111 2000 2000 0014
anne mots homolog Historia as Su attent for motise gen - October 11, 2002 of any accessional first of accessional firs	rinnary fest (1 et have assigned the h MC - restore flax (match regarding the match regarding the match regarding the cost, Med. Cost allow cost, Med. Cost allow cost Nandos involve a jen (Cost allow et hat Nandos involve a jen (Cost allow et hat Nandos in Tebu (Cost a	<ul> <li>Polbeling annotations to the gene product of MSEC (The text reflects annotations at characters at the comments indiction bits exist at the state exist 1944 (1917) for a state is a model and the bits exist at the state exist 1944 (1917) for an exist present of a state three formation of a state in the state exist 1944 (1917) for an exist at the period of a state in the state exist 1990 (1917) for an exist of the state exist 1990 (1917) for an exist at the state three formations are the state exist 1990 (1917) for a state exist at the state exist 1990 (1917) for an exist of the state exist at the state exist 1990 (1917) for an exist at the state exist 1990 (1917) for a state exist 1990 (1917) for a state exist at the state exist 1990 (1917) for a state exist at the state exist 1990 (1917) for a state exist 1990 (1910) for a sta</li></ul>	10200 2000 10111 2000 10555 10555
anne mots homolog Historia 1991) annotations as Su accreter House ger - Colorer House ger - Colorer House ger - Colorer House ger annotations as Su annotations as Su Subaran Molecu Functo annotations as Su Nolecu Functo accolorer Molecu Functo accolorer Molecu Functo Molecu Functo Molecu Functo Molecu Functo Molecu Functo Molecu Functo Molecu Functo Molecu Functo	rimary lest (1 et have assigned the h Mill relation (1) (match relation (1) (match relation (1) (match relation (1)) (match re	<ul> <li>Polowing annotations to the gene product of MahC (The text reflects annotations as of inskelpenes, meillered to meter induct for each alterative as if MarC (1012) for a structure as inference on meter induction each alterative as if MarC (1012) for a structure as inference on the text of the structure as if MarC (1012) for a structure as inference on the structure and we as and we as an inference on the structure with VGSI MSC. NGSI MSC inference on the structure as infere</li></ul>	10200 2000 10111 2000 10555 10555
anne mots homolog Historia 1991) annotations as Su accreter House get - Colorer House get - Colorer House get - Colorer House get annotations as Su annotations as Su Subara annotations as Subara Subara annotations as Subara Suba	rinnary fest (1 esthave assigned the h MC - reation of this mation regarding the mation regarding the mation regarding the cost, Med Coletter Ve a per OC et al. Autors a per OC et al. Autors	<ul> <li>Pole-ing an obstitues to the gene a vocation MEPC (The test villetes an obstitues as of using pere summelian of the fill residue theory of MEPC (The test villetes an obstitues as of MePC 2017) for a set of pereoperative of the know. Fileste support test set of MePC 2017 for a set of test set of the know. Fileste support test set of MEPC 2017 for a biological residue with VCSI MSD, VCSI MSD, and VCSI MSD, and VCSI MSD, and VCSI MSD and VCSI MSD are set of the view of the view of test set of the VCSI MSD, and VCSI MSD, and VCSI MSD are set of the view of test set of the VCSI MSD, and VCSI MSD, and VCSI MSD are set of the view of test set of the VCSI MSD, and VCSI MSD are set of the view of test set of the VCSI MSD, and VCSI MSD are set of the view of test set of the VCSI MSD, and VCSI MSD are set of the view of test set of the VCSI MSD, and VCSI MSD are set of the view of test set of the VCSI MSD. And VCSI MSD are set of the view of test set of the VCSI MSD. And VCSI MSD are set of the view of test set of the VCSI MSD and VCSI MSD are set of the view of test set of the VCSI MSD. And VCSI MSD are set of the VCSI MSD are set of the VCSI MSD are set of the view of test set of test set of the VCSI MSD. And VCSI MSD are set of the V</li></ul>	2000 P 2000 P 2000 P 2000 P 2000 P 2000 P
anne mots homolog Historia 1991) annotations as Su accreter House get - Coloner House get - Coloner House get - Coloner House get annotations as Su annotations as Su Sologitht annotations as Su Subarantees and Molecu Functor annotations as Su Nolecu Functor accologitht accolo	rimary lest () esthave assigned the h MC - restore () has mation regarding the mation regarding the mation regarding the a stat Mach C ref at Ma- a that Mach C ref at Ma- a that Mach C ref at Ma- a that Mach C ref at Ma- a part () at all success motations in Table () Constitution () Constitution (	<ul> <li>Probleming an notations to the gene production WHC (The test reflects an notations at children as uncelland of the firm end at the strict set of Merch 2012. Then, is not set on a training produces of the kinet fractions and the firm end at the strict set of Merch 2012. Then, is not set on a strict kinet in the stellar KOD SC (101114) KOD SC (10121114) and a NuoMed ID.</li> <li>Product set of the strict head of the firm the stellar KOD SC (101114) KOD SC (10121114) and a NuoMed ID.</li> <li>Product set of the strict head of the stellar KOD SC (101114) KOD SC (10121114) and the strict set of the stellar test set on the total set of Merch 2012 of the stellar set o</li></ul>	2000 P 2000 P 2000 P 2000 P 2000 P 2000 P 2000 P
anne mots homolog Historia 1991) annotations as Su accorden House get a Coloner House get annotations as Su annotations as Su Subarant Modesu Functo annotations as Su Subarant Modesu Functo accordens a Modesu Functo reprodus Nolesu Functo	rinnary fest (1 esthave assigned the h Mill restore flax (match regarding the match regarding the match regarding the a cal, Mad. Chef all Ya- a that Nand Shef all Ya- (Chesan all Sheft) (Chesan all Ya- a that Nand Sheft) (Chesan all Ya- All Ya- (Chesan all Ya- (Chesan all Ya- All Ya- (Chesan all Ya- (Chesan	<ul> <li>Problem gramssetters to the gene a vector MetC (The test veloces a most or site of the set of the dimensional of the time is a threat relax of MetC 2017. The set of the problem of the time is a threat relax of MetC 2017. The set of the set of the time is a threat relax of MetC 2017. The set of the problem of the time is a threat relax of MetC 2017. The set of the set of the time is a threat relax of MetC 2017. The set of the set of the time is a threat relax of MetC 2017. The set of the set of the time is a threat relax of MetC 2017. The set of t</li></ul>	10000           10000
anne mots homolog Hereit ions as Su accreter Hereitse ger - Coloren	rimary lest () esthave assigned the h Mill reation of this mation regarding the mation regarding the mation regarding the align () at all sources restances in Table () Constitution align () at all sources () Constitution () Con	Problem gramsset ers to the gene avecut or MatC (The test veloces an operators as of invergeness, mellena connecter of the thin edia broad one as if Med 2010 2016 and set is a set segme to be been as interesting the delete KGI NC of PET 14 KGI NC 0 CEED 11 and a DuXNed ID.          Important and the period of a diversity the delete KGI NC of PET 14 KGI NC 0 CEED 11 and a DuXNed ID.         Important and the period of a diversity the delete KGI NC of PET 14 KGI NC 0 CEED 11 and a DuXNed ID.         Important and the period of a diversity the delete KGI NC of PET 14 KGI NC 0 CEED 11 and a DuXNed ID.         Important and the period of a diversity the delete KGI NC of PET 14 KGI NC 0 CEED 11 and a DuXNed ID.         Important and the period of a diversity the delete KGI NC of PET 14 KGI NC 0 CEED 11 and a DuXNed ID.         Important Network (Text View)       (OO Oraph)         Important (Text View)       (OO Oraph)	10000           10000
Innet Schemolog         Hotel Constant         State of the method set of the set	rimary lest () esthave assigned the h Mill reation () fins (mation regarding the mation regarding the mation regarding the cold, Med. Chet all va- cold, Med. Chet a	<ul> <li>Polo-inglamostions to the gene a vector WHC (This test vibers annotations as of invergences, methods an interferend in the fine edge travelet as a MARC (3017) for a strends gene pression and the fine edge gene critical and a NuMed ID.</li> <li>Problem and the problem and the fine edge gene critical and a NuMed ID.</li> <li>Problem and the problem and the fine edge gene critical and a NuMed ID.</li> <li>Problem and the problem and the problem and the fine edge gene critical and a NuMed ID.</li> <li>Problem and the problem and th</li></ul>	<ul> <li>b. 500</li> <li>b. 500</li></ul>
anne mots homolog Hereit ions as Sur acceler Hereit Parts and any accertana into anne fuel and We si 100510 ouggeste anne fuel and We si 200510 ouggeste Melan mel and dat anne fuel si 200510 ouggeste Nelaon health anne fuel si 200510 ouggeste Nelaon health anne fuel si 200510 ouggeste Nelaon health anne fuel si 200510 ouggeste Nelaon health anne fuel si 200510 Nelaon health Nelaon health Nelaon health Nelaon health Nelaon health Nelaon health Nelaon health Nelaon health Nelaon health Nelaon health Nelaon health	rimary lest () esthave assigned the h Mill reation of this mation regarding the mation regarding the mation regarding the end, Med. Cheffel We a per Direct Alexandes rest Kishbis involve a compact CNA bis fine rest Kishbis involve rest Kishbis	<ul> <li>Problems anosotions to the gene a could be WebC (The test reflects an assotions ad of outer press, method and an additional in each at the set of Need 2011 / 1 and set of a call the problems and the time each of test and a flavor test press, and test and test and a flavor test press, and test and te</li></ul>	1:000           1:000
Image       Image       Image       Image         Image       Hotel I (1997)         Image       H	rinnary fest (1 est have assigned the h Millimetarian offers (mation regarding the mation regarding the mation regarding the a cal, Madil Chef all Ma- a that Manb Sherve ver- a jam (Classification) (mations in Tabu (mations in Tabu (ma	<ul> <li>Toballing an acctor size the gene proceed of XEPC (The best follows a most or size of a most or size the gene process of the those flows and the time size gene process of the those flows flows flows and the time size gene process of the those flows flows</li></ul>	<ul> <li>b.253-</li> <li>b.253-</li></ul>
Innet Schemolog         Hard Schemolog         State Schemolog<	rimary lest () esthave assigned the h Mill reation () fins mation regarding the mation regarding the mation regarding the algen () at all success algen () at all success motations in Table () Constitution algen () at all success () Constitution algen () at algen () () Constitution () Consti	<ul> <li>Probleming annocation size the gene proceed of WHC (The best follows a needed of a line of the first second of the fi</li></ul>	<ul> <li>b.2000</li> <li>b.2000</li></ul>



#### **Curation-driven Ontology Improvement**

**Coordinated curation leads to modification and improvement of the ontology.** For example, the new term "guanine/thymine mispair binding" was created for reference genome annotation because experimental data showed that the human MSH6 gene product, as well as the orthologous genes in mouse, zebrafish and yeast, are involved in this function.

SF.net » Projects » Gene Ontology » Tracker » Curator requests » Browse Tracker	r Items			
Gene Ontology		🔍 Trackers	Search Advanced	
Project - Tracker - Mailing Lists - Forums - Code - Services	🕶 Download 🔹 Documenta	ion Tasks	Project Web Site 🗹	
Login and Submit New Browse Admin			🖄 Stats  🕅 RSS	
Assignee: (?) Status: (?) Category: (?)	Group: (?) Reference Genomes			
	,			
Show only: Submitter username : Summary keyw				
Sort By: (?) ID    Descending	Browse			
Browse the term requests currently under consideration. Request ID Summary	Open Date	Priority Status Assigned	To Submitted By	
Trequest ID Summary	•	Thomy Status Assigned	To Subinitied By	
1541912 NTR: children of mismatched DNA binding	* 2006-08-17 05:15	6 Closed gomidori 🗡	vkhodiyar	
DouiGR: Term Deballa - Mazilla I				<u>_ic</u>
	rain Tone Joh			
±the o	Gene Ontology		AmiC	GO
	Accessional Res	ron BL49T scarch Brown	Trees	
5.x	anda dia <mark>100,7070,87</mark>	R Trims – Gröberg on violans – 🗖 Exa	ert Matuh <mark>Buomt S. cry</mark>	
	guanine/	thymine mispair bin	ding	
		" Treage & Become Concrete & Terr	acsociatione *	
Term Informatio	n			

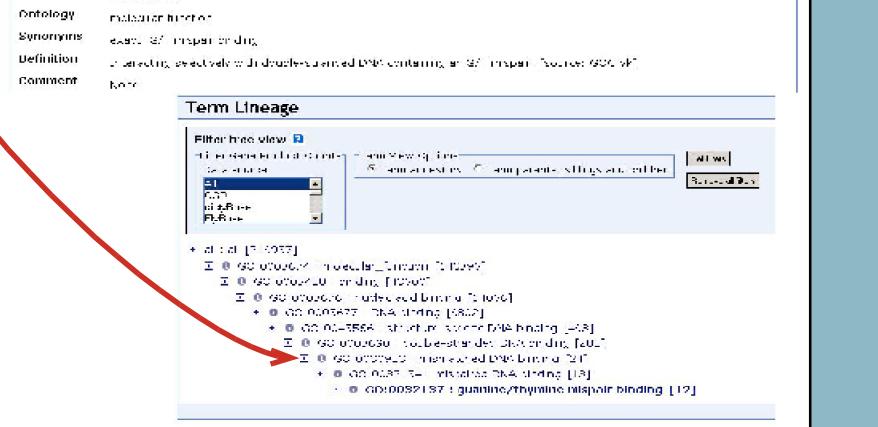
Danio rerio	<i>MSNO</i>
Arabidopsis thaliana	MSH6, MSH6-1
Saccharaomyces cerevisiae	MSH6
Schizosaccharomyces pombe	msh6
Escherichia coli	mutS

ACHE	ATXN10	DPYS	MSH2	PIP5K3
ACR	ATXN2	DRD2	MSH3	PML
ACTC1	ATXN3	DRD3 🤇	MSH6	POLA1
ACVR1	ATXN7	DRD4	MYBPC3	PPT1
ACVR1B	B4GALN2	DRD5	MYH11	PPT1
ACVR2A	B4GALT1	EIFAK3	MYH6	PRKAG2
ACVRB2	BAX	EIF2B1	MYH7	PSEN1
ACVRL1	BCL10	EIF2B2	MYH8	PSEN2
ADSL	BCL2	EIF2B3	MYH9	RECQL4
AGGF1	BCL3	EIF2B4	MYL2	RMRP
AGXT	BCL6	EIF2B5	MYL3	RPS19
ALAS2	BCL7A	ERCC6	MYLK2	SBDS
ALB	BCL8	ERCC8	MY01A	SETX
ALDH5A1	BLM	FOXC1	MY03A	SHROOM3
ALG12	BRCA2	FOXC2	MY05A	SHROOM4

#### Many genes are chosen that are associated with human disease

S NCBI	Online	OMI] Mendelian Inheri	M itance in Man	Johns Hopkins University	,			My NCBI E (Sign In) (Register)
All Databases	PubMed	Nucleotide	Protein	Gename	Structure	PMC	Taxonomy	ONIM
Search OMIM	v for 2600678		Go Clea	a Save Search				
	Limits Preview/Index	History Clipboa	rd Details					
Entrez	Display Detailed	Y Show 20	Y Sendto Y					
OMM Recent OMM	All: 31 OMIM UniSTS: 5	OMIM dbSNP: 5	k					
Search OMIM Search Gene Map		Items 1 - 20 of 3	1				P	age 1 of 2 New
Search Morbid Map	□1: <u>+600678</u>							GeneTests, Link

**MGI GO** annotations are coordinated with other homolog annotations for the selected gene to provide an overview of the gene's function.



**Our goal to achieve comprehensive annotation** is an effort to capture all the unique features of a gene product that can be characterized in the **Gene Ontology (GO). Detailed comprehensive** annotation coordinated across species presents an overview of a gene that a biologist would recognize as describing that particular gene.

Alternative	tities; symbols	
	ICH-BINDING PROTEIN; GTBP AL CANCER, HEREDITARY NONPOL	YPOSIS, TYPE 5, INCLUDED; HNPCC5, INCLUDE
TABLE OF (	ONTENTS	
MAPF     MOLE     PHEN     ANIM     ALLEI     CONT     CREA	FUNCTION	



padopoulos et al. (1995) found a 1-bp deletion mutation of the MSH6 gene in the HCT-15 colorectal cancer cell line at codon 222, which changed a leucine to a termination codon. They also found a 5-bp deletion/substitution at codon 1103 (TTGATAGAGT to TTTGT), which created a new termination codon 9-bp downstream. 😞

0002 REMOVED FROM DATABAS

DMIM Help

How to Link

Numbering System

ymbols How to Print

Citing OMIM

Download

OMIM Facts

Statistics Jpdate Log estrictions on U

Allied Resources

Genetic Alliance

Databases

Locus-Specific

Model Organism

enotype

MitoMap

HGMD

0003 REMOVED FROM DATABA

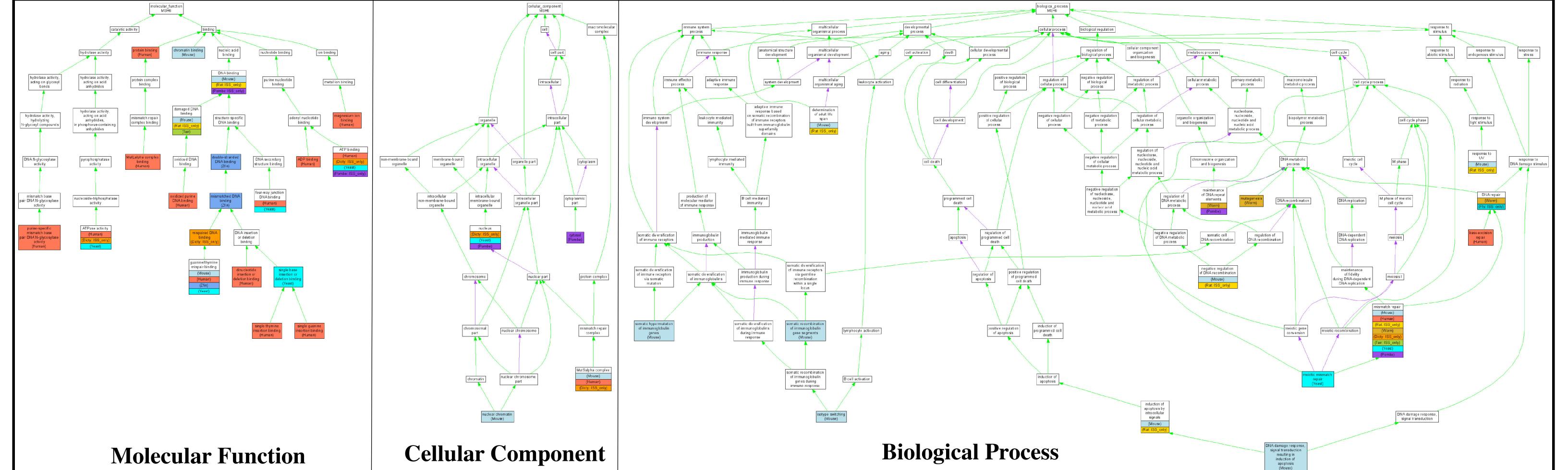
#### 0004 COLORECTAL CANCER, HEREDITARY NONPOLYPOSIS, TYPE 5 [MSH6, 1-BP DEL, F\$570TER

fiyaki et al. (1997) demonstrated an MHS6 germline mutation in an HNPCC family in which no mutation could be identified in the MSH2 (60 (436) gene. Carcinoma of the transverse colon in a 52-year-old member of the family showed alterations in 5 of 7 dinucleotide repeat loci and all 4 heterozydosity at 50, 8p, 17p, and 18g, or mutations of the TP53 (191170) and KRAS2 (190070) denes. At age 53 this patient also had an endometri: carcinoma, which exhibited RER(+) at 3 of 4 dinucleotide repeats and 3 of 4 mononucleotide repeats. PCR-SSCP analysis of DNA from the colon and endometri carcinoma and normal tissues detected a mutant band for MSH6. Direct sequencing of the mutant band revealed oredicting premature stop at codon 570 and truncation of the MSH6 protein. The same permline mutation was also detected in the patient's sister, who presente mutation, as it was detected in their offspring. In addition to the germline mutation, somatic mutations of MSH6 were detected in carcinomas from the proband ii codons. These somatic mutations were presumably in the alleles without the germline mutation, suggesting that inactivation of both alleles of MSH6 was the caus of the RER(+) phenotype and the stimulus for neoplasia. Although this family did not fulfill the full 'Amsterdam criteria,' path endometrial, ovarian, and pancreatic carcinomas. Mivaki et al. (1997) considered it noteworthy that endometrial and ovarian carcinomas were predominant in thi amily, in contrast to the predominance of colorectal carcinomas in families with MSH2 or MLH1 cermline mutations. The mean age for carcinoma formation in this family was 58 years, which is somewhat later than the mean age of 41 years for the first appearance of cancer in the usual HNPCC families with germline mutation of MSH2 or MLH1. 🥥

**OMIM records can provide information about the** human disease.

**Graphical view of annotations for human** *MSH6* **and its orthologs** 

molecular_fun
MOUC



Graphical views provide the "big picture" of how genes function in all model organisms. In this view, different colors in the boxes represent annotations from different model organisms.

# http://www.informatics.jax.org

## http://www.geneontology.org