### LITERATURE TRIAGE AND INDEXING IN THE MOUSE GENOME INFORMATICS (MGI) GROUP

Randal P. Babiuk, Dale A. Begley, Susan M. Bello, Dirck W. Bradt, Donna L. Burkart, Howard Dene, Alexander D. Diehl, Harold J. Drabkin, Jacqueline H. Finger, Terry F. Hayamizu, David P. Hill, Michelle Knowlton, Debra M. Krupke, Ira Lu, Lois J. Maltais, Monica McAndrews-Hill, Terrence F. Meehan, Li Ni, Hiroaki Onda, Dmitry Sitnikov, Constance M. Smith, Cynthia L. Smith, and Monika Tomczuk. The Jackson Laboratory, Bar Harbor, Maine, USA 04609

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

Preventing genomic instability

### **ABSTRACT**

The Mouse Genome Informatics (MGI; http://www.informatics.jax.org) group is comprised of several collaborating projects including the Mouse Genome Database (MGD) Project, the Gene Expression Database (GXD) Project, the Mouse Tumor Biology (MTB) Database Project, and the Gene Ontology (GO) Project. Literature identification and collection is performed cooperatively amongst the groups.

In recent years many institutional libraries have transitioned from a focus largely on print holdings to one of electronic access to journals. This change has necessitated adaptation on the part of the MGI curatorial group. Whereas the majority of journals covered by the group used to be surveyed in paper form, those journals are now surveyed electronically. Approximately 160 journals have been identified as those most relevant to the various database groups. Each curator in the group has the responsibility of scanning several journals for articles relevant to any of the database projects. Articles chosen via this process are marked as to their potential significance for various projects. Each article is catalogued in a Master Bibliography section of the MGI database system and annotated to the database sections for which it has been identified as relevant. A secondary triage process allows curators from each group to scan the chosen articles and mark ones desired for their project if such annotation has been missed on the initial scan.

Once articles have been identified for each database project a variety of processes are implemented to further categorized and index data from those articles. For example, the Alleles and Phenotype section of the MGD database indexes each article marked for MGD and in this indexing process they identify each mouse gene and allele examined in the article. The GXD database indexing process has a different focus. In this case articles are indexed with regard to the stage of development used in the study as well as the assay technique used. In each case the indexing gives an overview of the data held in the article and assists in the more extensive curation performed in the following step of the curation process. Indexing also provides each group with valuable information used to prioritize and streamline the overall curation process.

## ~160 Journals Reviewed



Cancer

MCB

e

Curation Statu

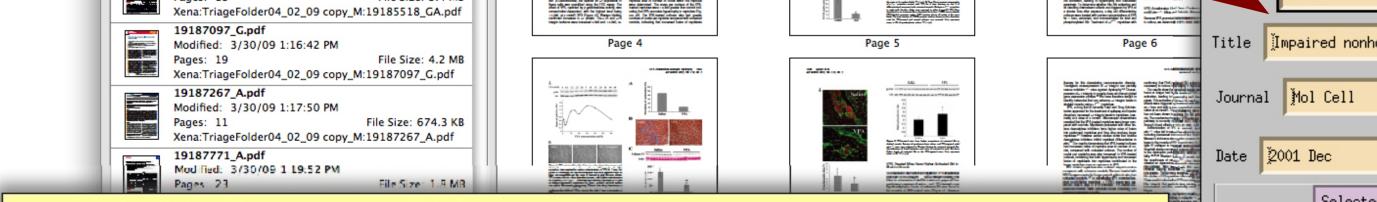
100 🔜 🖬 📰 Browse

Research

The MGI projects are supported by NHGRI grants HG000330, HG002273, HG003622, NICHD grant HD033745, and NCI grant CA089713.

Image: Strain of the open mark of the open	QUOSA         Image: Control in the phosphatic dyise mer-laceled tumor cells by cend?         Published       Surce	Development
<ul> <li>Image: Constraint of the second second</li></ul>	y-drug conjugates for th butions, myelosuppressi -inducible factor-lalpha inhibition of Stat3 indu mediated down-regulati in of constitutively activa ate increases expression of it attivates	to MGI per year
Relevant articles identified and QUOSA for various database	Subscience       Subscience<	File       Commands       Reports       NLM         Type       Article       In       NLM?       Yes         Review       Status       Peer       Reviewed       Is       Review Article?       No         rs       Bharpless NE; Ferguson D0; O'Hagan RC; Castrillon DH; Lee C; Farazi PA; Alson S; Fleming J; Mor       Title       Impaired nonhomologous end-joining provokes soft tissue sarcomas harboring chromosomal transloc
Database Areas:	Pages: 19 Kena:TriageFolder04_02_09 copy_M:19187097_G.pdf 19187267_A.pdf Modified: 3/30/09 1:17:50 PM Pages: 11 Kena:TriageFolder04_02_09 copy_M:19187267_A.pdf 19187771_A.pdf	Journal Mol Cell
Alleles & Phenotype	PDF files collected in the primary journal scan are collected	Selected Used Not Used Never Used Selected Never Used
Gene Expression	into folders and one curator from each database area scans	Probes/Seq     X     Tumor     X       Mapping     X     SCC
Gene Ontology (GO)	the articles to identify any missed for their area in step 1.	Allele/Phe       X       X       Image: Allele/Phe       X       X         Homology       Image: Allele/Phe       X       Image: Allelee/Phe       X
Mapping	Select All 77 files	Expression       X       X       Image: Second se
Nomenclature	Acrobat used to further identify database areas	Nomen       X       tations are cataloged.         Search Data Sets Using: ◇ AND ◆ OR (default)       CR (default)

Tumor



Master Bibliography catalogs all articles and their chosen area associations

### Gene & Allele Indexing Mammalian glycerol uptake/transpo known as Skn, the homolog of the Drosoph mammalian Gup1 has a O-acyltransferase family. residue in the motif that is indispense

# **Gene Expression Indexing**

	-	-	-	-				-		
?			(	Gene		<b>pres</b> ry Res			Literature	
<u>You searched for</u> Authors: contains <i>Elstrott</i> Journal: equals <i>Neuron</i> Year: equals 2008						-			Each relevant reference is	
34 matching records from 1 refere	nce								indexed for gene expression	
Summary by Age and Ass	<b>ау:</b> Nı	Imbers	in the	table ii	ndicate	the nu	ımber o	of re	data, cataloged by gene, ag	
Age	E11.5	E12.5	E13.5	E14.5	E15.5	E16.5	E17.5	A	and assay used.	
In situ protein (section)	<u>1</u>		<u>1</u>	2	<u>1</u>	<u>6</u>	2	4	and assay used.	
In situ RNA (section)				<u>3</u>		20				
In situ protein (whole mount)		1		<u>3</u>				1		

# **Tumor Indexing**

Browse	MTB Index Database	
Layout: DataEntry 🕨	New Paper Search List View Indexing	Home
+ + +	J:Number J:125183 O Journal Nat Med Year 2007	Organ(s) mammary gland 2) lymphohematopoletic 3) skin 4) liver
Record: 16 Found: 452	Priority Top	5) lung 6) cardiovascular 7) soft tissues
Total: 10088 Sorted		8) 9) 10)
		11) 12)

(Abe Y, et al., FEBS J 2008 Jan;275(2):318-31)

that was fused to

Each relevant reference is indexed for the genes and alleles used and standard nomenclature is applied.

### Experimental Procedures

Mouse Colony and Histopathology lig4 <sup>+/-</sup> ink4a/arf <sup>+/-</sup> mice were produced by a cross of lig4 <sup>+/-</sup> males	
with ink4a/arf <sup>-/-</sup> fem described (Frank et a were ina mixed 129/S the effects of backgro posed of lithermate animals. Animals were checked for tumors three	p
Lig4 tmlFwa overall survival between bice was also statistically	Γ
significant (p = 0.007, log rank test). For the determination of radio-	
(Sharpless, et al., Mol Cell 2001 Dec;8(6):1187-96)	

Western blot				<u>4</u>
RT-PCR		<u>3</u>		

Summary

Actb actin

Actg1 acti **Results Re** 

Ntrk1 neu

Immary by Gene and R	eference: Number indicates the number of results matching the search criteria recorded for each reference.
actin, beta       Results       Reference       1       J:145295       Elstrott       established       independent	?       Gene Expression Literature         Query Results Details         Symbol       Ntrk1
ctg1actin, gamma, cytopiResultsReference1J:145295establishedindepende	Symbol       neurotrophic tyrosine kinase, receptor, type 1         ID       MGI:97383         Reference       J:145295       Elstrott J; Anishchenko A; Greschner M; Sher A; Litke AM; Chichilnisky EJ; Feller MB, "Direction selectivity in the retina is established independent of visual experience and cholinergic retinal waves." Neuron 2008 May 22;58(4):499-506
Itrk1       neurotrophic tyrosing         Results       Reference         4       J:145295         Elstrott J;       established independe	<ul> <li>Indicates gene expression was analyzed but not necessarily detected.</li> <li>Age</li> <li>E14.5</li> <li>E16.5</li> <li>A</li> <li>In situ protein (section)</li> <li>In situ RNA (section)</li> <li>In situ protein (whole mount)</li> <li>In situ RNA (whole mount)</li> <li>In situ reporter (knock in)</li> </ul>
	In situ reporter (knock in) Northern blot Western blot RT-PCR CDNA clones RNase protection Nuclease S1

Authors Zhang L; Anglesio MS; O'sullivan M; Zhang F; Yang G; Sarao R; Nghiem MP; Cronin S; Hara H; Melnyk N; Li L; Wada T Liu PP; Farrar J; Arceci RJ; Sorensen PH; Penninger JM

Abstract Transformation and cancer growth are regulated by the coordinate actions of oncogenes and tumor suppressors. Here, we show that the novel E3 ubiquitin ligase HACE1 is frequently downregulated in human tumors and maps to a region of chromosome 6q21 implicated in multiple human cancers. Genetic inactivation of HACE1 in mice results in the development of spontaneous, late-onset cancer. A second hit from either environmental triggers or genetic heterozygosity of another tumor suppressor, p53, markedly increased tumor incidence in a Hace1-deficient background Re-expression of HACE1 in human tumor cells directly abrogates in vitro and in vivo tumor growth, whereas downregulation of HACE1 via siRNA allows non-tumorigenic human cells to form tumors in vivo. Mechanistically, the tumor-suppressor function of HACE1 is dependent on its E3 ligase activity and HACE1 controls adhesion-dependent growth and cell cycle progression during cell stress through degradation of cyclin D1. Thus, HACE1 is a candidate chromosome 6q21 tumor-suppressor gene involved in multiple cancers.

> PubMed ID 17694067 Date Indexed October 24, 20 Each relevant reference Date Coded October 29, Coded\_By dmk is indexed for the organs studied for tumorigenesis and is also prioritized for data entry.