Ira A. Fulton School of Engineering
ARIZONA STATE UNIVERSITY
School of Computing and Informatics

Relationship Extraction from Documents

Chitta Baral
Professor
Arizona State University

Why?

- Goal: Use text mining and automated reasoning to help research in sciences
- The particular domain area we have focused on: Molecular Biology

Questions of our interest – at a high level

- Semantic Search of documents
- Explain (a set of) observations; make a diagnosis based on observations
- Predict the impact of particular interventions
- Design a drug therapy.
- Generate hypothesis regarding hitherto unknown aspects of a bio-process.

Text Mining: two aspects

- Extract facts from text
  - Automatics Extraction
  - Collaborative development of databases
- Obtain more general knowledge from the text

Extracting Facts from Text

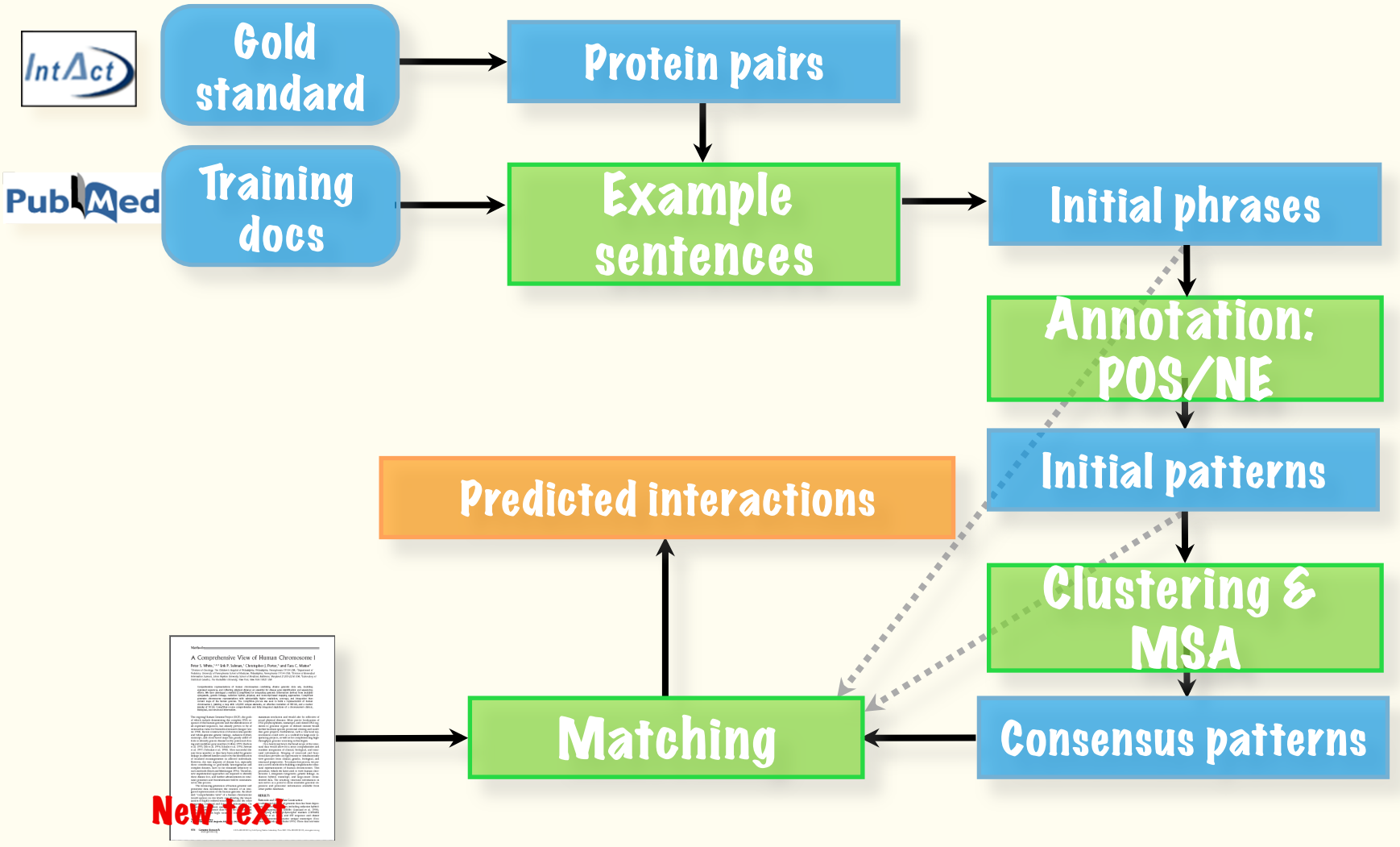
- For example, some of the azole antifungals are inhibitors of both P450 enzymes and P-glycoprotein (Nivoix et al., 2008), whereas rifampicin is an inducer of both CYP3A4 and P-glycoprotein (Katragadda et al., 2005).

Extracting more general knowledge from text

- While the importance of metabolism in many drug-drug interactions is beyond question, it has become increasingly apparent in recent years that inducers and inhibitors of some of the enzymes of drug metabolism can also affect drug transporter proteins.
- For example, some of the azole antifungals are inhibitors of both P450 enzymes and P-glycoprotein (Nivoix et al., 2008), whereas rifampicin is an inducer of both CYP3A4 and P-glycoprotein (Katragadda et al., 2005). (page 2)
- Hence, interaction can sometimes involve drug-metabolizing enzymes, drug transporters, or both.

Extracting facts from text: protein-protein interactions

Yappie – Work flow



Yappie – Initial phrases

- >120,000 snippets that discuss PPI, such as

.. P suppressed P ..
.. P helps regulate P ..
.. P recruits the adapter molecule P ..
.. P binding domain of P ..
.. P binds directly to the extracellular domain of P ..
.. P associates with a novel P dependent kinase, P ..
.. while P activation reduces P expression/activation ..
.. P was previously found to interact with the KRAB silencing domain of P and with the P ..

Yappie - Phrase alignment

Initial phrase 1 →

Initial phrase 2 ↓

		DT	PTN	VB	PRP	ADJ	PTN	CC	PTN
	0	0	0	0	0	0	0	0	0
PTN	0	0	4	0	0	0	4	0	4
VB	0	0	0	7	6	0	0	2	0
PRP	0	0	0	6	8	7	0	0	0
PTN	0	0	4	5	6	5	11	5	4

Initial phrase 1: DT PTN VB PRP ADJ PTN CC PTN

Initial phrase 2: PTN VB PRP - PTN

Yappie – Multiple phrase alignment

Initial phrases:

protein	strongly	binds	to		protein
protein		interacts	with	the	protein
protein	never	binds	to		protein
protein		regulates		the	protein
protein		inhibits		a	protein

Consensus pattern:

PROTEIN	{strongly,never}	{binds, .., ..}	{to, with}	{the, a}	PROTEIN

would exactly match the sentence (part):

protein		binds	to	the	protein

Performance: PPI extraction

- #4 system in **BioCreative 2** for protein-protein interactions (**2007**)
- f-measure of 24%, respectively (1st: 30%)
- 20 participants

- #1 system for PPIs in **BioCreative II.5 (2009)**
- 30% f-score (2nd: 23%)
- 15 participants
- >100 submissions overall (multiple configurations per participating team allowed)

- Main Person leading this at ASU: Joerg Hakenberg

BioCreative II.5 challenge

- Participated 2 of 3 tasks
  - INT: Interactor normalization task (1st)
  - IPT: Interaction pair task (1st)
- http://www.biocreative.org/news/chapter/biocreative-ii5/
- Main person in our group on this: Joerg Hakenberg

SNPshot of PubMed

SNPshot: Aim

- collect information on genes regarding
  - genetic variants / mutations / alleles,
  - associations with diseases,
  - drug interactions (transport, metabolism; activation, inhibition),
  - allele frequencies and populations
- large-scale, fully automated
- from Medline abstracts
- link to evidence and cross-link to other databases for validation and further information

Entities & relations

- genes & proteins
- drugs
- diseases
- genetic variants / mutations
- SNPs / alleles / haplotypes
- populations & frequencies

Gene - drug
Gene - disease
Drug - disease
Gene - variant
Gene - allele
Gene - RefSNP
Frequency - allele
Frequency - variant
Frequency - population
Variant - population
Allele - population

- MutationFinder [CBR+07], 700 regular expressions; added 100 more
- BANNER [LG08@PSB]

Normalization

- map genes, drugs, diseases to database identifiers (EntrezGene, Uniprot; PharmGKB, DrugBank; UMLS)
- canonical form for variants (HGVS: c.76A>T)
- map SNPs to RefSNP/dbSNP
- populations to canonical form
- plain dictionary matching for drugs & diseases
- GNAT for genes & proteins [HPL+08]
- heuristics for all others

Data sets

- PubMed abstracts
- PharmGKB: 3614 referenced PubMed citations
- 40 VIP PGx genes from PharmGKB
- expanded using PubMed's "Related Articles" functionality ➠ 26,000 additional abstracts
- PubMed query ➠ 30,000 abstracts

- around 58,000 abstracts

```
(phenotype OR haplotype OR genotype OR mutation OR allele OR variant OR SNP OR polymorphism)
AND (disease OR risk OR disorder OR malfunction) AND (drug OR pharmalogical OR metabolize
OR inhibit OR bioavailability OR orally) AND human[MH] AND hasAbstract,
```

Relationship extraction

- mostly simple heuristics
- sentence-level co-occurrence + keywords (for different kinds of relations: [CKY+08])

Method	Total relations per set 3,614	55,095
Co-occ with keywords	14,968	186,983
Simple co-occ	2,195	37,949
Respectively	97	637
LCA sub-tree	795	5,166
Co-occ for 1:n	3,101	24,085
Co-occ for m:n	1,524	9,314
Low confidence co-occ	1,790	11,588
Total	24,470	275,722

Generalizing text extraction: Querying Parse Trees

Motivation

- Traditional information extraction technique works as a pipeline
  - Perform grammar parsing, named entity identifier, named entity recognizer, normalization, extraction
- Information extraction is seen as a one-time process
- Common issues in the development of extraction system
  - What if we change our extraction goals?
    - e.g. extract gene-disease associations rather than protein-protein interactions
  - What if we have an improved NER system?
  - Which of the extraction patterns work well?

Motivation

- Information extraction should *not* be seen as a pipeline or one-time process
- With the pipeline approach, need to re-extract from the entire text collection
  - Computationally expensive!
- But change of extraction goals or improvement of components *does not affect* the entire text collection
  - if we extract gene-disease associations, only need to extract from sentences that have gene and disease mentions
  - if we deploy a new NER, only sentences that are newly tagged are needed to perform re-extraction

What's needed for extraction?

- To minimize reprocessing, we need to store parse trees and semantic information
  - a database is ideal to store information that we need to perform extraction
- Extraction should be seen as generic
- Can we use database queries as information extraction?
  - Hard to express syntactic patterns with SQL
  - We need a new query language for extraction, called parse tree query language (PTQL)

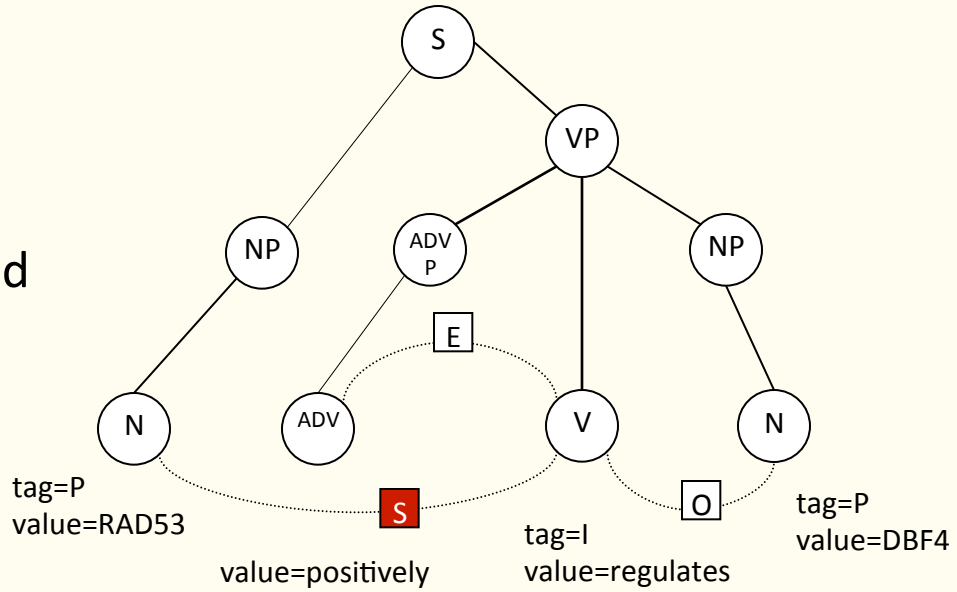
Parse trees

- Stores dependency linkages and constituent trees
- Linkage: shows the dependencies between words in a sentence

```
   +---------S---------+
   |         +----E-----+---O----+
   |         |          |        |
RADB53 positively regulates.v DBF4 .
```

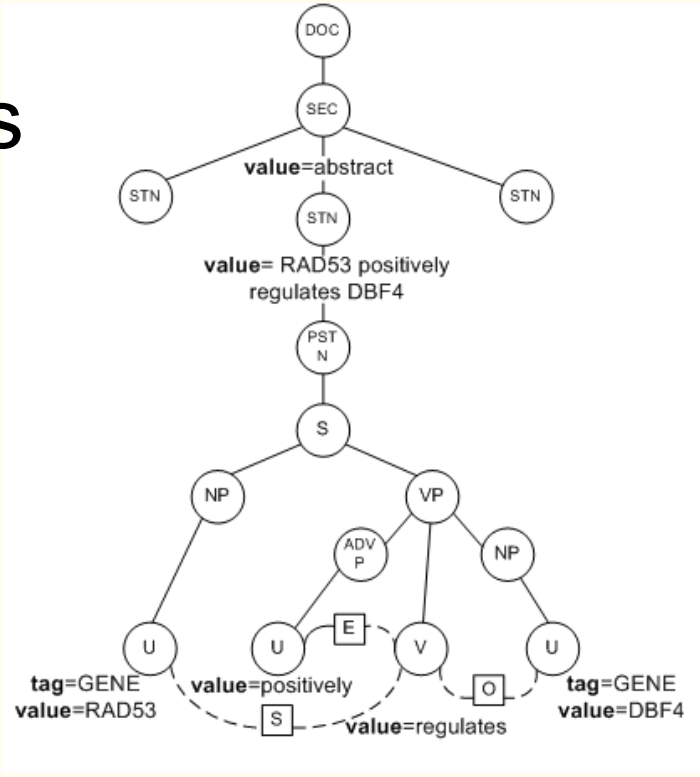
- S: connects subject-noun
- E: verb-modifying adverbs
- O: transitive verbs to direct or indirect objects

- Constituent trees are represented “vertically”
- Linkages are represented “horizontally”



Parse Tree Database

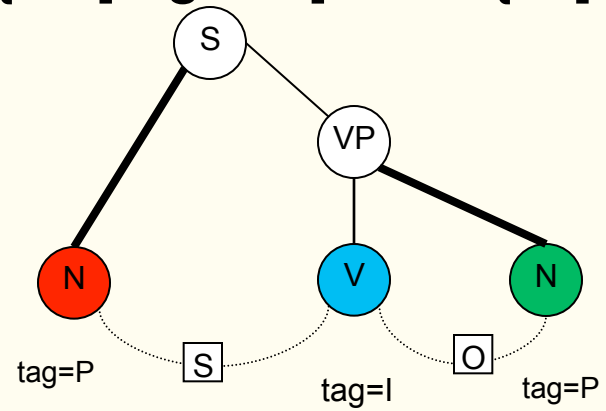
- Represents a document with its sentences and parse trees in a hierarchy
- Uses a labeling scheme
- Certain important properties:
  - Given a parse tree, for any pair of nodes *q* and *p*,



  - *q* is a child of *p* iff *q*.pid = *p*.id
  - *q* is a descendant of *p* iff *q*.left ≥ *p*.left, *q*.right ≤ *p*.right and *q*.depth > *p*.depth
  - *q* immediately follows *p* iff the left most child of *q* immediately follows the right most child of *p*, i.e., q.left = p.right
  - *q* follows *p* iff *q.left* ≥ p.right

PTQL query syntax

- A PTQL query has 4 components in this format
  - *tree pattern* : *link condition* : *proximity condition* : *return expression*
- Tree pattern
  - X{...Y...}: Y is a node in the subtree with X as the root
  - /: parent/child relation in the constituent tree
  - //: ancestor/descendant relation in the constituent tree
    - Example: //S{//N[tag='P']->/VP{/V[tag='I']->//N[tag='P']}}



Other applications of PTQL

- **Feature extraction**
  - Find all MeSH terms and their frequencies among documents that contain recognized gene names.
    - //DOC(x) { //?[tag='GENE'] } : : : count(x.mesh), x.mesh
- **Normalize gene names**
  - Find articles x of some author in which gene *y* is mentioned.
    - //DOC(x)[author='John Smith']{//?[tag='GENE'](y)}::: distinct x.value, y.value
- **Normalize gene names to species**
  - Find gene-species relations based on some grammatical patterns, such as gene and species occurring in the same noun phrase.
    - //S{//NP{//N[value='human']=>//?[tag='GENE'](x)}} ::: x.value
- **Boosting recall for gene name recognizer**
  - Suppose *"p53"* has been tagged as a gene name in some documents, find *"p53"* such that *"p53"* is not tagged as a gene name.
    - //DOC(x){//STN(y){//?[tag!='GENE' and value='p53']}}::: x.value, y.value

- **Can we have a system/language that utilizes syntactic dependencies for IR and IE tasks?**

- **Our solution**: integrate IR and a parse tree database with a query language called IR+PTQL
  - database stores parse trees (with syntactic dependencies)
  - allows IR and parse tree database (PTQL) to work independent of each other
- **Goals**: Flexible use of IR and database
  - application-independent
  - keep both subsystems (IR and parse tree database) "untouched" with the use of a *middleware*
  - utilize syntactic dependencies in IR
  - improve efficiency for information extraction (IE)

IR+PTQL by example


```
1 +glycan* +modif* >>
    //STN{//?[value LIKE {'%glycan%','%modif%'}](w1)<=>
2   //N[value LIKE {'%glycan%','%modif%'}](x)->//?[value='i.e.']=>//N(y)}::
      w1.value, x.value, y.value >> [alldoc] +w1.value +(x.value y.value^0.5) 3
```

1. Retrieve documents using the IR query (subquery 1)
2. Among the documents retrieved by subquery 1, find values of w1, x and y that satisfy the syntactic pattern
3. Using the values from subquery 2, form IR query in subquery 3 by enhancing subquery 1

Examples of applications: building pathways

Building pathways

- An important part of understanding or reverse-engineering biological phenomena (disease, phenotype, etc.)
- Connecting the dots !!!
- Building pathways involves
  - Connecting the dots, where the dots are
    - Biological data (such as interactions)
  - But an equally important aspect is
    - Biological Knowledge and
    - Reasoning with that knowledge

More examples: studying drug-drug interactions

Importance of studying drug-drug interactions

- **Drug design:** Early assessment of a new compound's potential interactions with other drugs can avoid costly investment in the drug discovery process.

- **Drug prescription:** For multi-drug prescription, pharmacokinetic interactions amongst co-administrated drugs may alter the bioavailability of the drugs that can lead to life-threatening side effects for the patients.

Looking beyond automatic extraction and manual curation of facts

Main Issues

- Manual curation is expensive
- Automatic curation still has many errors
- What to do?

Key Idea

- If lots of articles are being written then lot of people are writing them and lot of people are reading them.
- If only we could make these people (the authors and the readers) contribute to the curation effort …
- Especially the readers; the ones who need the curated data!

Mass collaboration has worked in

- Wikipedia
- Project Gutenberg
- Netflix rating
- Amazon rating
- Etc.

Mass collaborative curation: initial hurdles

- An average reader
  - (S)he is not normally interested in filling a blank curation form.
  - We can not make an average reader go though curation training.
  - So it has to be very different from just making the existing curation tools available to the mass and expect them to contribute.

Mass collaborative curation : key initial ideas

- Make it very easy:
  - user need not remember where (which database, which web page) to put the curated knowledge.
  - Curation opportunity should present itself seamlessly.
- Curation should not be a burden to an average user
  - Make the curated knowledge “thin”.
- There should be immediate rewards
  - Do not start with a blank slate.

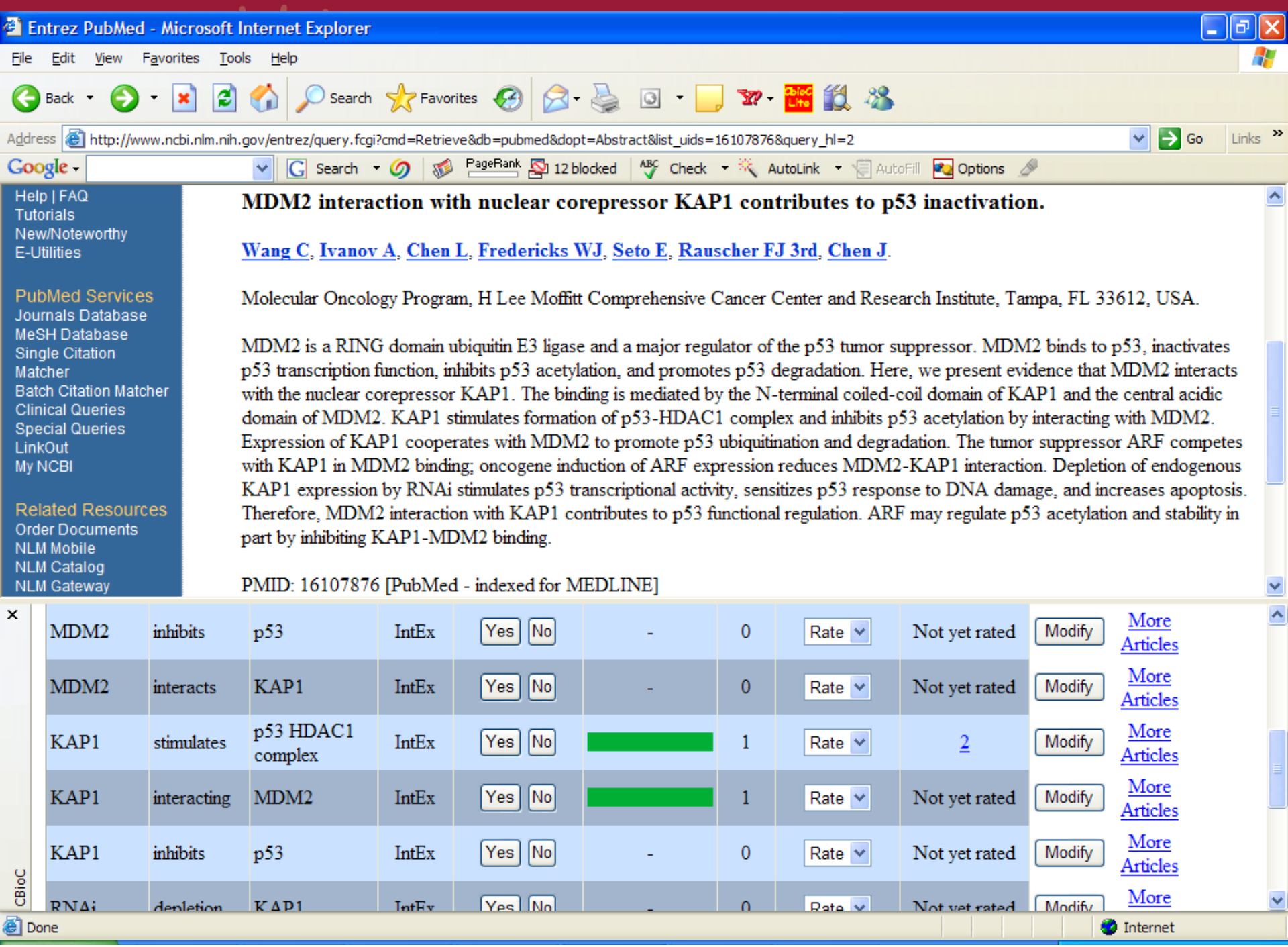
MDM2 interaction with nuclear corepressor KAP1 contributes to p53 inactivation.

Wang C, Ivanov A, Chen L, Fredericks WJ, Seto E, Rauscher FJ 3rd, Chen J.

Molecular Oncology Program, H Lee Moffitt Comprehensive Cancer Center and Research Institute, Tampa, FL 33612, USA.

MDM2 is a RING domain ubiquitin E3 ligase and a major regulator of the p53 tumor suppressor. MDM2 binds to p53, inactivates p53 transcription function, inhibits p53 acetylation, and promotes p53 degradation. Here, we present evidence that MDM2 interacts with the nuclear corepressor KAP1. The binding is mediated by the N-terminal coiled-coil domain of KAP1 and the central acidic domain of MDM2. KAP1 stimulates formation of p53-HDAC1 complex and inhibits p53 acetylation by interacting with MDM2. Expression of KAP1 cooperates with MDM2 to promote p53 ubiquitination and degradation. The tumor suppressor ARF competes with KAP1 in MDM2 binding; oncogene induction of ARF expression reduces MDM2-KAP1 interaction. Depletion of endogenous KAP1 expression by RNAi stimulates p53 transcriptional activity, sensitizes p53 response to DNA damage, and increases apoptosis. Therefore, MDM2 interaction with KAP1 contributes to p53 functional regulation. ARF may regulate p53 acetylation and stability in part by inhibiting KAP1-MDM2 binding.

PMID: 16107876 [PubMed - indexed for MEDLINE]


Entrez PubMed - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites
Address http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=pubmed&dopt=Abstract&list_uids=16107876&query_hl=2
Go Links
Google Search PageRank 12 blocked Check AutoLink AutoFill Options
Help | FAQ
Tutorials
New/Noteworthy
E-Utilities
PubMed Services
Journals Database
MeSH Database
Single Citation Matcher
Batch Citation Matcher
Clinical Queries
Special Queries
LinkOut
My NCBI
Related Resources
Order Documents
NLM Mobile
NLM Catalog
NLM Gateway
MDM2 interaction with nuclear corepressor KAP1 contributes to p53 inactivation.
Wang C, Ivanov A, Chen L, Fredericks WJ, Seto E, Rauscher FJ 3rd, Chen J.
Molecular Oncology Program, H Lee Moffitt Comprehensive Cancer Center and Research Institute, Tampa, FL 33612, USA.
MDM2 is a RING domain ubiquitin E3 ligase and a major regulator of the p53 tumor suppressor. MDM2 binds to p53, inactivates p53 transcription function, inhibits p53 acetylation, and promotes p53 degradation. Here, we present evidence that MDM2 interacts with the nuclear corepressor KAP1. The binding is mediated by the N-terminal coiled-coil domain of KAP1 and the central acidic domain of MDM2. KAP1 stimulates formation of p53-HDAC1 complex and inhibits p53 acetylation by interacting with MDM2. Expression of KAP1 cooperates with MDM2 to promote p53 ubiquitination and degradation. The tumor suppressor ARF competes with KAP1 in MDM2 binding; oncogene induction of ARF expression reduces MDM2-KAP1 interaction. Depletion of endogenous KAP1 expression by RNAi stimulates p53 transcriptional activity, sensitizes p53 response to DNA damage, and increases apoptosis. Therefore, MDM2 interaction with KAP1 contributes to p53 functional regulation. ARF may regulate p53 acetylation and stability in part by inhibiting KAP1-MDM2 binding.
PMID: 16107876 [PubMed - indexed for MEDLINE]
CBioC
MDM2 inhibits p53 IntEx Yes No - 0 Rate Not yet rated Modify More Articles
MDM2 interacts KAP1 IntEx Yes No - 0 Rate Not yet rated Modify More Articles
KAP1 stimulates p53 HDAC1 complex IntEx Yes No 1 Rate 2 Modify More Articles
KAP1 interacting MDM2 IntEx Yes No 1 Rate Not yet rated Modify More Articles
KAP1 inhibits p53 IntEx Yes No - 0 Rate Not yet rated Modify More Articles
RNAi depletion KAP1 IntEx Yes No - 0 Rate Not yet rated Modify More
Done
Internet

CBioC Summary

- Information/curation window pops up automatically.
- Automatic extraction is used as a boot strap so that no user is working on a blank slate.
- Users vote on correctness, make corrections, add fact.
  - Suppose 60% precision and recall of automatic extraction system
  - A person will have an easier time discarding 40% of wrongly extracted text than identifying 60% of correct entries and entering them!

Collaborative Bio Curations 2

CBioC2 p53 2001[edat] Search History | Settings | Logout

Introduction: The volume of existing biomedical articles is huge and it grows day by day. From 1994 to 2004, close to 3 million biomedical articles were published by US and European researchers alone. Added to the approximately 15 million abstracts already in PubMed, this represents over 800 new articles per day and a myriad of individual new facts to survey for information relevant to a particular research question.

Currently two approaches are pursued to extract and combine facts from biomedical publications. The first approach of hiring human curators is expensive, and thus does not scale-up. It also leads to bias. The second approach of using automated information extraction systems only has a recall and precision of around 60%.

We present here a new approach to the problem through mass collaboration, where the community of researchers that writes and reads the biomedical texts will be able to contribute to the curation process, dictating the pace at which it is done.

How you can help: Automated text extraction is used as a starting point to bootstrap the database, but then we need help from you and your team of biologists to improve upon the extracted data and help 'iron out' inconsistencies on a massive scale.

Use the search bar to find articles in the PubMed database. Open the document and view both human and algorithmic curations. You can then add your own curations or you can vote on how correct our algorithms or users are.

CBioC2 p53 2001[edat] Search Group | History | Settings | Login

Authors	Publisher	Pub. Date	pmid
Katano H, Sato Y, Sata T.	Cancer	2001Dec	11753987
Expression of p53 and human herpesvirus-8 (HHV-8)-encoded latency-associated nuclear antigen with inhibition of apoptosis in HHV-8-associated malignancies.			
Benassi M S, Molendini L, Gamberi G, Magagnoli G, Ragazzini P, Gobbi G A, Sangiorgi L, Pazzaglia L, Asp	Cancer	2001Dec	11753985
Involvement of INK4A gene products in the pathogenesis and development of human osteosarcoma.			
Esposito I, Friess H, BÃ¼chler M W.	Zentralblatt fur Chirurgie	2001Nov	11753795
Molecular mechanisms in chronic pancreatitis.			
Jiang X H, Wong B C, Lin M C, Zhu G H, Kung H F, Jiang S H, Yang D, Lam S K.	Oncogene	2001Nov	11753684
Functional p53 is required for triptolide-induced apoptosis and AP-1 and nuclear factor-kappaB activation in gastric cancer cells.			
Zhang D, Vuocolo S, Masciullo V, Sava T, Giordano A, Soprano D R, Soprano K J.	Oncogene	2001Nov	11753676
Cell cycle genes as targets of retinoid induced ovarian tumor cell growth suppression.			
MÃ¼nger K, Basile J R, Duensing S, Eichten A, Gonzalez S L, Grace M, Zacny V L.	Oncogene	2001Nov	11753671
Biological activities and molecular targets of the human papillomavirus E7 oncoprotein.			
Mantovani F, Banks L.	Oncogene	2001Nov	11753670

View | Page 1 of 137 | Abstracts | Search by relation | Displaying articles 1 - 25 of 3422

Interaction: Protein Protein

Bio Entity 1	Interaction	Bio Entity 2	Articles	Collection
beta-catenin		E-cadherin	25	1561
NF-kappaB		AP-1	15	228
wnt-1		int-2	3	938
NK-kappaB		AP-1	14	1946
wnt-1		int-2	11	3081
wild-type KLF6		p21	24	1976
Bax		Bcl-2	20	1511
p53		p21	11	2199
biclonal anticytokeratin antibody		AE1	20	3045

CBioC2 p53 2001[edat] Search Group | History | Settings | Login

Authors	Publisher	Pub. Date	pmid
Katano H, Sato Y, Sata T.	Cancer	2001Dec	11753987

Expression of p53 and human herpesvirus-8 (HHV-8)-encoded latency-associated nuclear antigen with inhibition of apoptosis in HHV-8-associated malignancies.

BACKGROUND: Kaposi sarcoma (KS) and primary effusion lymphoma (PEL) cells express human herpesvirus-8 (HHV-8)-encoded latency-associated nuclear antigen (LANA) (open reading frame [ORF] 73 protein), suggesting that LANA plays an important role in the pathogenesis of HHV-8-associated malignancies. Recently, the binding of LANA to p53 was demonstrated in vitro. In the current study, the authors investigated the association between p53 and LANA expression with apoptosis in HHV-8-associated malignancies in vivo. METHODS: Twenty-six cases of KS, 1 case of HHV-8-associated solid lymphoma, 2 PEL cell lines, and an HHV-8-associated lymphoma engrafted in severe combined immunodeficiency (SCID) mice were examined. Immunohistochemistry using the catalyzed signal amplification system was employed to detect LANA and p53 on paraffin embedded tissues and the immunofluorescence technique was used on cell lines. To detect apoptosis, the TdT-mediated dUTP nick end labeling (TUNEL) method was used. For mutation analysis of p53, exons 5-9 of the p53 gene were amplified by polymerase chain reaction and examined by direct sequencing. RESULTS: Immunohistochemistry revealed that LANA and p53 were expressed in the tumor cells of all these specimens, and apoptotic cells were rarely detected in them using the TUNEL method. Immunofluorescence assay revealed that LANA colocalized with p53 in the nuclei of PEL cells. Sequencing analysis indicated that there was no mutation in the deduced amino acid sequences of p53 in KS tissues. CONCLUSIONS: These data suggest colocalization of p53 and LANA and the inhibition of apoptosis in HHV-8-associated malignancies in vivo, supporting the results found in vitro that p53 inhibition by LANA suppresses cell death, as reported previously. These results also suggest that the p53 pathway is crucial in the pathogenesis of HHV-8-associated malignancies.

Benassi M S, Molendini L, Gamberi G, Magagnoli G, Ragazzini P, Gobbi G A, Sangiorgi L, Pazzaglia L, Asp	Cancer	2001Dec	11753985

Involvement of INK4A gene products in the pathogenesis and development of human osteosarcoma.

BACKGROUND: The INK4A tumor suppressor gene plays a crucial role in the regulation of the G1 cell cycle phase. It encodes two transcripts, p16 and p14 alternate reading frame (ARF), involved in retinoblastoma protein (pRb)- and p53- cell growth control pathways, respectively. METHODS: To define the role of gene status and molecule expression involved in the INK4A regulatory system, immunohistochemistry, immunoblotting, and polymerase chain reaction (PCR) analysis were performed on 35 primary high grade osteosarcomas (OS). RESULTS: Although p16 and p14ARF proteins were found negative or weakly detectable in 60% and 57% of the cases respectively, INK4A gene analysis of exons 1alpha, 1beta and 2 did not reveal any deletion or mutation. However, methylation status of the 5'CpG promoter region, assessed by methylation-specific PCR, was found in 12 out of 21 OSs with negative or weak p16 expression. A statistical analysis based on pRb/p16 and p53/p14ARF staining status showed that pRb and p16 co-expression was inversely correlated to tumor relapse and was a marker for a more favorable prognosis. A statistically significant inverse correlation was found between wt-p53 and p14ARF expression. In the group of wt-p53 tumors, the loss of p14ARF was associated with a decreased expression of p21 protein, suggesting a down-regulation of the transcriptional activity of p53. CONCLUSIONS: The current results suggest that, in OS, the altered expression of INK4A products plays a primary role in the deregulation of both pRb and p53 cell growth control pathways, contributing to tumor pathogenesis and development.

Esposito I, Friess H, BÃ¼chler M W.	Zentralblatt fur Chirurgie	2001Nov	11753795

Molecular mechanisms in chronic pancreatitis.

The pathogenesis of chronic pancreatitis (CP) is still controversial. None of the proposed models has been able to provide a convincing link between the known etiological factors - alcohol abuse, metabolic disturbances, congenital or acquired obstruction of the duct system - and the complex

View | Page 1 of 137 | Abstracts | Search by relation | Displaying articles 1 - 25 of 3422

Interaction: Protein Protein

Bio Entity 1	Interaction	Bio Entity 2	Articles	Collection
beta-catenin		E-cadherin	25	1561
NF-kappaB		AP-1	15	228
wnt-1		int-2	3	938
NK-kappaB		AP-1	14	1946
wnt-1		int-2	11	3081
wild-type KLF6		p21	24	1976
Bax		Bcl-2	20	1511
p53		p21	11	2199
biclonal anticytokeratin antibody		AE1	20	3045

Wnt-1 and int-2 mammary oncogene effects on the beta-catenin pathway in immortalized mouse mammary epithelial cells are not sufficient for tumorigenesis.

Hollmann C A, Kittrell F S, Medina D, Butel J S.

Oncogene

Development of strategies for prevention of breast cancer development requires an understanding of the effects of mammary oncogenes on mammary cells at early stages in neoplastic transformation. As mammary oncogenes wnt-1 and int-2 affect different signal transduction pathways, we investigated their effects on established mouse mammary epithelial cell lines (MMECLs) reflecting early stages in tumorigenesis. Normal interactions between beta-catenin and E-cadherin were abrogated in all three immortalized MMECLs and the cells lacked beta-catenin-mediated transactivation activity, detectable using a reporter assay, suggesting that alterations in cell adhesion may be very early events in mammary tumorigenesis. Immortalized FSK4 and EL12 cells and hyperplastic TM3 cells were stably transfected with expression vectors encoding wnt-1 or int-2 or the control vector, and drug-selected pooled cells from each line were confirmed by reverse transcription-polymerase chain reaction to express the transfected oncogene; this expression persisted in the cells analysed in vitro and in vivo. Resultant phenotypic changes depended both on the oncogene and the target mammary cell line. In FSK4 cells, expression of wnt-1 or int-2 resulted in proliferative changes in vitro, including reduced contact inhibition, increased beta-catenin expression, and decreased p53 transcriptional activity, but neither oncogene conferred upon those cells the ability to produce tumors in vivo. EL12 cells were highly refractory to the effects of both oncogenes, with the only measurable changes being increased E-cadherin levels induced by both oncogenes and increased proliferation of the int-2-transfected cells in the absence of serum. Parental TM3 cells were phenotypically similar to wnt-1- or int-2-transfected FSK4 cells and displayed an increased rate of proliferation in vitro and markedly increased tumorigenicity in vivo following transfection with int-2 but not with wnt-1. These results suggest that wnt-1 signaling is redundant in the hyperplastic TM3 cells and indicate that wnt-1-induced effects in the immortalized FSK4 and EL12 cells were not sufficient to mediate a tumorigenic phenotype. This study showed that the wnt-1 and int-2 oncogenes have similar but distinguishable effects on immortalized MMECLs and that the genetic background of the mammary cells greatly influences the consequences of oncogene expression at early stages of cell transformation.

PMID: 11753642

Add a relation | Upload Changes

Bio Entity 1	Relation or Entity Type	Bio Entity 2	Source	Collection	Agree/Disagree
Relation: Protein Protein					
beta-catenin	Unknown	E-cadherin	Url	Url	1/0 vote
wnt-1	Unknown	int-2	Url	Url	1/0 vote
wnt-1	Unknown	int-2	Url	Url	0/1 vote

Wnt-1 and int-2 mammary oncogene effects on the beta-catenin pathway in immortalized mouse mammary epithelial cells are not sufficient for tumorigenesis.

Hollmann C A, Kittrell F S, Medina D, Butel J S.

Oncogene

Development of strategies for prevention of breast cancer development requires an understanding of the effects of mammary oncogenes on mammary cells at early stages in neoplastic transformation. As mammary oncogenes wnt-1 and int-2 affect different signal transduction pathways, we investigated their effects on established mouse mammary epithelial cell lines (MMECLs) reflecting early stages in tumorigenesis. Normal interactions between beta-catenin and E-cadherin were abrogated in all three immortalized MMECLs and the cells lacked beta-catenin-mediated transactivation activity, detectable using a reporter assay, suggesting that alterations in cell adhesion may be very early events in mammary tumorigenesis. Immortalized FSK4 and EL12 cells and hyperplastic TM3 cells were stably transfected with expression vectors encoding wnt-1 or int-2 or the control vector, and drug-selected pooled cells from each line were confirmed by reverse transcription-polymerase chain reaction to express the transfected oncogene; this expression persisted in the cells analysed in vitro and in vivo. Resultant phenotypic changes depended both on the oncogene and the target mammary cell line. In FSK4 cells, expression of wnt-1 or int-2 resulted in proliferative changes in vitro, including reduced contact inhibition, increased beta-catenin expression, and decreased p53 transcriptional activity, but neither oncogene conferred upon these cells the ability to produce tumors in vivo. EL12 cells were highly refractory to the effects of both oncogenes, with the only measurable changes being s and oncogenes and increased proliferation of the int-2-transfected cells in the absence of serum. Parental TM3 cells were phenotypically displayed an increased rate of proliferation in vitro and markedly increased tumorigenicity in vivo following transfection with int-2 but t-1 signaling is redundant in the hyperplastic TM3 cells and indicate that wnt-1-induced effects in the immortalized FSK4 and EL12 cells phenotype. This study showed that the wnt-1 and int-2 oncogenes have similar but distinguishable effects on immortalized MMECLs and that the genetic background of the mammary cells greatly influences the consequences of oncogene expression at early stages of cell transformation.

PMID: 11753642

View | Recent Activity

- [x] Protein/Gene
- [x] Drug
- [x] Disease
- [x] Species
- [x] Bioprocess
- [x] Cell
- [x] Cellular Component
- [x] Multiple
- [x] Relation

wnt-1: Defined as Protein/Gene
Search wnt-1 as Protein/Gene:10 or 20 articles
Source: BANNER and/or MetaMap

Add a relation | Upload Changes

Bio Entity 1	Relation or Entity Type	Bio Entity 2	Source	Collection	Agree/Disagree
Relation: Protein Protein					
beta-catenin	Unknown	E-cadherin	Url	Url	1/0 vote
wnt-1	Unknown	int-2	Url	Url	1/0 vote
wnt-1	Unknown	int-2	Url	Url	0/1 vote

Wnt-1 and int-2 mammary oncogene effects on the beta-catenin pathway in immortalized mouse mammary epithelial cells are not sufficient for tumorigenesis.

Hollmann C A, Kittrell F S, Medina D, Butel J S.

Oncogene

Development of strategies for prevention of breast cancer development requires an understanding of the effects of mammary oncogenes on mammary cells at early stages in neoplastic transformation. As mammary oncogenes wnt-1 and int-2 affect different signal transduction pathways, we investigated their effects on established mouse mammary epithelial cell lines (MMECLs) reflecting early stages in tumorigenesis. Normal interactions between beta-catenin and E-cadherin were abrogated in all three immortalized MMECLs and the cells lacked beta-catenin-mediated transactivation activity, detectable using a reporter assay, suggesting that alterations in cell adhesion may be very early events in mammary tumorigenesis. Immortalized FSK4 and EL12 cells and hyperplastic TM3 cells were stably transfected with expression vectors encoding wnt-1 or int-2 or the control vector, and drug-selected pooled cells from each line were confirmed by reverse transcription-polymerase chain reaction to express the transfected oncogene; this expression persisted in the cells analysed in vitro and in vivo. Resultant phenotypic changes depended both on the oncogene and the target mammary cell line. In FSK4 cells, expression of wnt-1 or int-2 resulted in proliferative changes in vitro, including reduced contact inhibition, increased beta-catenin expression, and decreased p53 transcriptional activity, but neither oncogene conferred upon those cells the ability to produce tumors in vivo. EL12 cells were highly refractory to the effects of both oncogenes, with the only measurable changes being increased E-cadherin levels induced by both proliferation of the int-2-transfected cells in the absence of serum. Parental TM3 cells were phenotypically similar to wnt-1- or int-2-transfected FSK4 cells and e of proliferation in vitro and markedly increased tumorigenicity in vivo following transfection with int-2 but not with wnt-1. These results suggest that wnt-1 he hyperplastic TM3 cells and indicate that wnt-1-induced effects in the immortalized FSK4 and EL12 cells were not sufficient to mediate a tumorigenic wed that the wnt-1 and int-2 oncogenes have similar but distinguishable effects on immortalized MMECLs and that the genetic background of the mammary e consequences of oncogene expression at early stages of cell transformation.

View | Recent Activity

- [x] Protein/Gene
- [x] Drug
- [x] Disease
- [x] Species
- [x] Bioprocess
- [x] Cell
- [x] Cellular Component
- [x] Multiple
- [x] Relation

Relations
Protein Protein
Protein Disease
Protein Drug
Disease Drug
Protein Cellular Component
Protein Bioprocess
Protein Cell
Drug Bioprocess
Drug Cell
Cell Disease
Single Entity

Upload Changes

Bio Entity 1	Relation or Entity Type	Bio Entity 2	Source	Collection	Agree/Disagree
Relation: Protein Protein					
beta-catenin	Unknown	E-cadherin	Url	Url	1/0 vote
wnt-1	Unknown	int-2	Url	Url	1/0 vote
wnt-1	Unknown	int-2	Url	Url	0/1 vote

Wnt-1 and int-2 mammary oncogene effects on the beta-catenin pathway in immortalized mouse mammary epithelial cells are not sufficient for tumorigenesis.

Hollmann C A, Kittrell F S, Medina D, Butel J S.

Oncogene

Development of strategies for prevention of breast cancer development requires an understanding of the effects of mammary oncogenes on mammary cells at early stages in neoplastic transformation. As mammary oncogenes wnt-1 and int-2 affect different signal transduction pathways, we investigated their effects on established mouse mammary epithelial cell lines (MMECLs) reflecting early stages in tumorigenesis. Normal interactions between beta-catenin and E-cadherin were abrogated in all three immortalized MMECLs and the cells lacked beta-catenin-mediated transactivation activity, detectable using a reporter assay, suggesting that alterations in cell adhesion may be very early events in mammary tumorigenesis. Immortalized FSK4 and EL12 cells and hyperplastic TM3 cells were stably transfected with expression vectors encoding wnt-1 or int-2 or the control vector, and drug-selected pooled cells from each line were confirmed by reverse transcription-polymerase chain reaction to express the transfected oncogene; this expression persisted in the cells analysed in vitro and in vivo. Resultant phenotypic changes depended both on the oncogene and the target mammary cell line. In FSK4 cells, expression of wnt-1 or int-2 resulted in proliferative changes in vitro, including reduced contact inhibition, increased beta-catenin expression, and decreased p53 transcriptional activity, but neither oncogene conferred upon those cells the ability to produce tumors in vivo. EL12 cells were highly refractory to the effects of both oncogenes, with the only measurable changes being increased E-cadherin levels induced by both oncogenes and increased proliferation of the int-2-transfected cells in the absence of serum. Parental TM3 cells were phenotypically similar to wnt-1- or int-2-transfected FSK4 cells and displayed an increased rate of proliferation in vitro and markedly increased tumorigenicity in vivo following transfection with int-2 but not with wnt-1. These results suggest that wnt-1 signaling is redundant in the hyperplastic TM3 cells and indicate that wnt-1-induced effects in the immortalized FSK4 and EL12 cells were not sufficient to mediate a tumorigenic phenotype. This study showed that the wnt-1 and int-2 oncogenes have similar but distinguishable effects on immortalized MMECLs and that the genetic background of the mammary cells greatly influences the consequences of oncogene expression at early stages of cell transformation.

PMID: 11753642

View | Recent Activity

- [x] Protein/Gene
- [x] Drug
- [x] Disease
- [x] Species
- [x] Bioprocess
- [x] Cell
- [x] Cellular Component
- [x] Multiple
- [x] Relation

Add a relation | Upload Changes

Bio Entity 1	Relation or Entity Type	Bio Entity 2	Source	Collection	Agree/Disagree
Relation: Protein Protein					
beta-catenin	Unknown	E-cadherin	Url	Url	1/0 vote
wnt-1	Unknown	int-2	Url	Url	1/0 vote
wnt-1	Unknown	int-2	Url	Url	0/1 vote
p53	Add	Add	User	-	

Conclusion

- Relationship Extractions can be of big help in Science.
  - We focused on molecular biology and pharmacology
- Applications
  - Semantic Search
  - Help in reasoning (prediction, explanation, design)
  - Discovery

Acknowledgements

- My students and post-doctoral researchers
  - Especially, Luis Tari, Joerg Hakenberg, Graciela Gonzalez, Bob Leaman, Vo Nguyen, Barry Lumpkin
- Funding agencies
  - NSF
  - Science Foundation of Arizona
  - ASU
  - IARPA; ONR