

Trawling the bibliome

Wynand Alkema

NIZO Food Research

Radboud University Medical Center

Bibliomics : The study of the bibliome

Wiktionary
 ['wɪkʃənri] *n.*,
 a wiki-based Open
 Content dictionary

Entry [Discussion](#) [Citations](#)

bibliomics

English [\[edit\]](#)

Noun [\[edit\]](#)

bibliomics (*uncountable*)

- The study of **bibliomes**

bibliome

[Contents](#) [\[show\]](#)

English [\[edit\]](#)

Etymology [\[edit\]](#)

From **bibliography** and **-ome**, as **genome**.

Noun [\[edit\]](#)

bibliome (*uncountable*)

- (*biochemistry, genetics*) The complete set of **biological journal articles** and associated information.



Wikipedia has an article on:
[Bibliome](#)

Data Flow in Research Projects

Knowledge:
New products
Life cycle
extension
Intellectual
Property

Literature
review



Develop
hypothesis

Analysis
interpretation

Perform
experiments

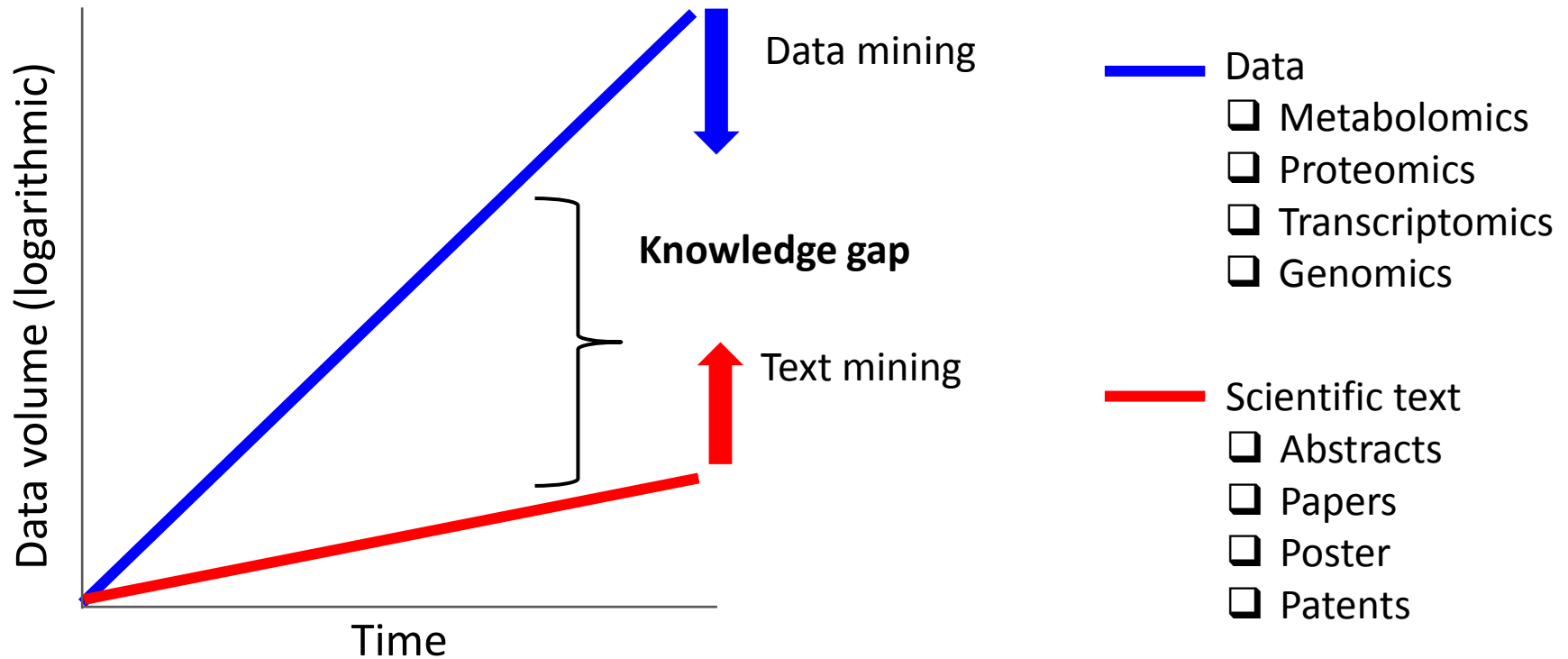


High Performance Computing



High Throughput Screening

Reduction of the knowledge gap



What is text mining?

The **computer assisted** screening of **large** volumes of text to **retrieve** relevant documents, **extract** useful information and **discover** new knowledge.

A relation in the text is a relation in real life

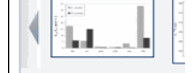
Display Settings: Abstract Send to: BMC Struct Biol. 2009 Mar 24;9:17. **Solution structure of the parvulin-type PPIase domain of Staphylococcus aureus PrsA- implications for the c**

Heikkinen O, Seppala R, Toss Department of Chemistry, Univers

Abstract
BACKGROUND: Staphylococ respiratory tract infections to lit antibiotics has created a need cytoplasmic lipoprotein which essential folding factor for sec S. aureus.
RESULTS: We have solved a h aureus PrsA (PrsA-PPIase). TI preference of the enzyme. With histidines we are able to give 1 dynamic behaviour of PrsA-PP
CONCLUSION: Detailed struct inhibitors. The structure resea of structural data is available c and our findings on the role of mechanism.

PMID: 19309529 [PubMed - indexe

Images from this publi



Publication Types, MeSH

LinkOut - more resource

United States Patent (117)

THE UNITED STATES OF AMERICA

United States - Wikipedia, the free encyclopedia - Microsoft Internet Explorer

http://en.wikipedia.org/wiki/United_States

United States

From Wikipedia, the free encyclopedia

This article is about the country in North America. For other uses, see United States (disambiguation) and United States (disambiguation).

The **United States of America**, also referred to as the **United States**, or simply the **U.S.**, and colloquially as **the States**, is a country in North America. It extends from the Atlantic Ocean to the Pacific Ocean and from the Gulf of Mexico to the Arctic Ocean. The United States is a federal republic, with 50 states, a federal district, and several territories. It is the world's third most populous nation, with a population of approximately 310 million as of 2010. The United States is a major world power, with a high Gross Domestic Product (GDP) and a strong military. It is also a major technological and cultural power, with a significant influence on the world.

PRODUCT X®

Brand name

Also called the active ingredient

Generic Name

Consumer Medicine Information

What is in this leaflet

Using Product X

Side effects

How to take it

After using it

What Product X is for

If you forget to take it

Storage

What to do if you miss a dose

Before you use Product X

When you must not use it

While you are using Product X

Things you must do

Product description

What Product X looks like

Ingredients

Check for possible interactions with your other medicines

Things to be careful of

Manufacturer

e.g. safe temperatures

What to do if you have leftover medicine

Also lists other ingredients or "fillers", i.e. what binds the active ingredients together

Contact details of the pharmaceutical company

Together to the next level

An explosion of tools

Name	Input	Output	Ref	Tasks	Description
askMEDLINE	Query	Abstracts	[17, 18]	IR	Free-text, natural language (English only) query for MEDLINE
Twease	Query	Ranked abstracts		IR	Indexes each word of Medline and provides features
XplorMed	Query	Ranked abstracts	[18]	IR	The system gives you the main associations between
eTBLAST	Tekst	Similar abstracts	[19]	IR	A text-similarity based search engine, using all words
MedlineRanker	Query	Ranked abstracts	[20]	IR	Ranks Medline abstracts based on user defined queries
MiSearch	Query	Ranked abstracts	[21]	IR	Ranks orders retrieved articles from PubMed based on
PICO	Query	Abstracts	[22]	IR	Search engine for MEDLINE/PubMed with Spelling Checker
PubCrawler	Query	Abstracts	[23]	IR	Free "alerting" service that scans daily updates of the
PubFocus	Query	Abstracts and S	[24]	IR	Statistical analysis of the MEDLINE/PubMed search
PubGet	Query	Abstracts + Link to PDFs		IR	A comprehensive source for science PDFs, including
Pubmatrix	List of Terms	Co-occurrence	[25]	IR	Allows simple text based mining of the NCBI literature
PubNet	List of Terms	Graph	[26]	IR	a flexible system for visualizing literature-derived ne
GeneValorization	GeneList	Abstracts and li	[27]	IR, NER	GeneValorization gives a very clear and handfull over
DPWP	GeneList	Linked concepts	[28]	IR, NER	Disease/ Phenotype PAGE is a disease focused gene s
Anne O'Tate	Query	Abstracts and C	[29]	IR, NER	Gives an overview of the set of articles retrieved by a
Chilibot	GeneList and Te	Graph	[30]	IR, NER	Chilibot searches PubMed literature database (abstr
ProteinCorral	Protein Name	Linked proteins	[31]	IR, NER	Combines Information Retrieval and Extraction from
MEDIE	Tripple	Highlighted phrases		IR, NER	Retrieve biomedical correlations from MEDLINE, bas
PubReminer	Query	Overrepresente	[32]	IR, NER	Breaks down a results of a PubmedQuery into categ
PubViz	Concepts	Hyperlinked gra	[33]	IR, NER	An Interactive Medline Search Engine Utilizing Extern
Quertle	Query + Concep	Abstracts + Ove	[34]	IR, NER	Using a combination of linguistic methods, Quertle f
Copub	Concept and Ge	Abstracts and li	[9]	IR, NER, IE	Web application with gene focussed retrieval of co-c
Coremine medical	Concept	Linked concepts		IR, NER, IE	Presents results about health, medicine and biology
Facta+	Tekst	Abstracts and li	[35]	IR, NER, IE	Finding Associated Concepts with Text Analysis
GoPubMed	Concept	Abstracts + Overrepresented coi		IR, NER, IE	Uses semantic algorithms to abstracts from the MED

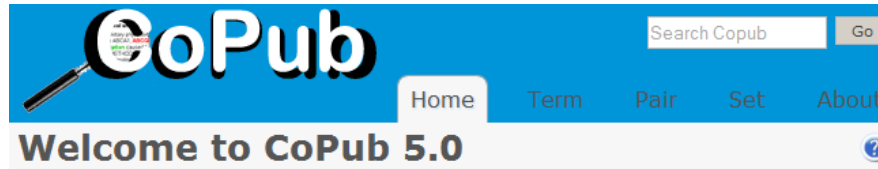
Mostly based on MEDLINE/ PubMed

Most of them co-occurrence based

Different tasks

- Document retrieval
- Information extraction
- Knowledge discovery

CoPub: a database of co-citations



CoPub is a text mining tool that detects co-occurring biomedical concepts in abstracts from the MedLine literature database.

*Nucleic Acids Research, 2011, 1-5
doi:10.1093/nar/gkr310*

CoPub update: CoPub 5.0 a text mining system to answer biological questions

Wilco W. M. Fleuren^{1,2*}, Stefan Verhoeven³, Raoul Frijters¹, Bart Heupers⁴,
Jan Polman³, René van Schaik³, Jacob de Vlieg^{1,3} and Wynand Alkema³

¹Computational Drug Discovery (CDD), CMBI, NCMLS, Radboud University Nijmegen Medical Centre,
²Netherlands Bioinformatics Centre (NBIC), PO Box 9101, 6500 HB Nijmegen, ³Molecular Design and
Informatics, MSD, PO Box 20, 5340 BH Oss and ⁴SARA Computing and Network Services, Amsterdam,
The Netherlands

Received January 27, 2011; Revised April 5, 2011; Accepted April 18, 2011

CoPub database

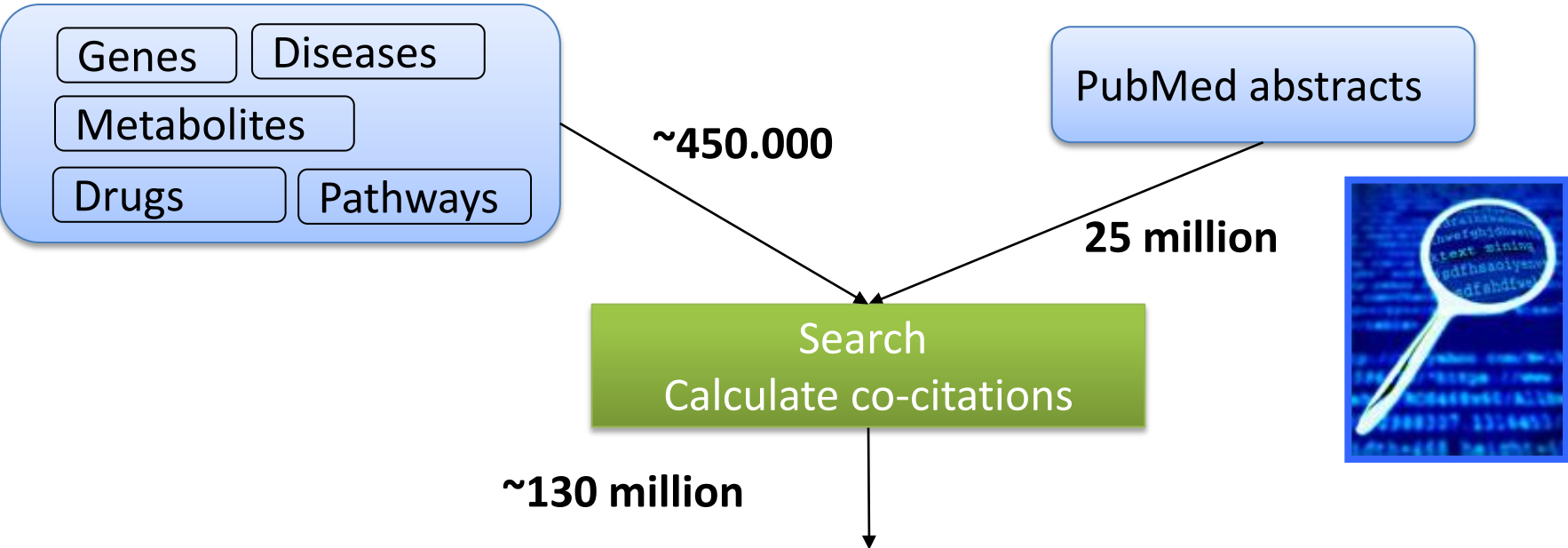
25 million papers

> 300 million literature relations

Statistical models for literature networks

www.copub.org

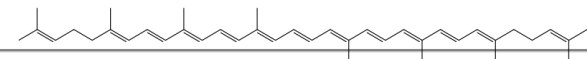
CoPub: A database of co-citations



Opposing effects of TGF-beta 2 on the Th1 cell development of naive CD4+ T cells isolated from different mouse strains.
 P Hoehn;S Goedert;T Germann;S Koelsch;S Jin;N Palm;E Ruede;E Schmitt. 1995-11-22. J Immunol. 155. PMID: 7561083

The development of naive dense CD4+ T cells from different mouse strains toward Th1 cells, as monitored by measuring secondary IFN-gamma production, was affected by TGF-beta 2 in a differential way. Th1 cell development of naive CD4+ T cells from strains C57Bl/6, BALB/c, and NMRI primed by immobilized anti-CD3 mAb was strongly inhibited in the presence of TGF-beta 2. Even when the Th1 cell-inducer IL-12 was added, the same effect of TGF-beta 2 was observed. In contrast, **Th1 development** was substantially promoted by TGF-beta 2 with T cells from C3H/He and CBA/J mice. Further analyses using CD4+ T cells from (C57Bl/6xCBA/J)F1 hybrids or DBA/1 mice showed that **Th1 development** was inhibited by TGF-beta 2 if the T cells were activated by anti-CD3 mAb, but it was enhanced upon costimulation with anti-CD28 mAb. Determination of primary **IL-2** production revealed that T cells from (C57Bl/6xCBA/J)F1 and DBA/1 mice produced low amounts of **IL-2** following stimulation by anti-CD3 mAb alone and comparatively high amounts after coactivation by anti-CD28 mAb. In the presence of TGF-beta 2, the production of **IL-2** was completely suppressed if such T cells were activated solely by anti-CD3 mAb, but it was only partially inhibited after costimulation by anti-CD28 mAb. Furthermore, TGF-beta 2-promoted **Th1 development** of such T cells was strongly inhibited after neutralization of endogenously produced **IL-2** and completely restored by the addition of human **IL-2**. Thus, our results indicate that the TGF-beta 2-mediated stimulation of Th1 cell development requires the presence of relatively high concentrations of **IL-2**. Therefore, the opposing effect of TGF-beta 2 on the Th1 cell development of naive CD4+ T cells from different mouse strains appears to be the result of the variable potency of the respective CD4+ T cells to produce **IL-2** in the presence of TGF-beta 2.

What is known about Lycopene?



GoPub Home Term Pa

Term details

Preferred name	lycopene
Alternative names	(6e,8e,10e,12e,14e,16e,18e,20e,22e,24e,26e)-2,6,10,14,19,23,27,31-octamethyldotriaconta-2,6,8,10,12,14,16,18,20,22,24,26e all-trans-lycopene psi,psi-carotene
Categories	metabolite
Nr of abstracts	2297

Actions:

Retrieve abstracts in which term occurs | Add term to set | Search for indirect relations with term

Below are the most significant co-occurrences of 'lycopene' with each category. (Click in the count column to show the abstracts in which terms co-occur)

R scaled score >= 30 | Abstract count >= 3 | Order by R scaled score

disease

Cancer of prostate | Eye diseases | Gastrointestinal neoplasm | Infertility, male | Maculopathy, age-related | Myoma | Vitamin a deficiency | cardiovascular disease | colon cancer | mouth neoplasm

metabolite

7,7',8,8',11,11',12,12'-octahydro-psi,psi-carotene | alpha-carotene | beta-zeacarotene | cyclic carotene | epsilon,epsilon-carotene | gamma-carotene | lutein | neurosporene | zeaxanthin | zeta-carotene

polyphenols

5-caffeoylquinic acid | curcumin | ellagic acid | gallic acid | kaempferol | myricetin | naringenin | naringin | quercetin | 3-o-rutinoside | resveratrol

GO molecular function

antioxidant | dioxygenase | glutathione peroxidase | glutathione reductase | hydratase | insulin like growth factor binding | monoxygenase | retinoic acid receptor | scavenger receptor | synthase

celltype

Adenocarcinoma cell | Breast cancer cell | Caco-2 cell | Cancer cell | Peripheral blood mononuclear cell | Premalignant cell | foam cell | peripheral blood lymphocyte | red blood cell

human gene

BXDC2 | CEL | GGT3 | GJA1 | IGFBP3 | JTV1 gene | PON1 | PTGDS | SCARB1 | XRCC1

tissue

abdominal adipose tissue | colon mucosa | endosperm | epithelial cancer | hypocotyl | mammary cancer | mycelia | prostate | sepal | uterine leiomyoma

drug

ascorbic acid | capsicum | dutasteride | finasteride | probucol | retinol | acetate | retinol palmitate | tocopherol acetate | vitamin a | vitamin c

pathway

apoptosis | cell communication | cell cycle | dna repair | g1 phase | immune system | lipid metabolism | m phase | mevalonate pathway | s phase

GO biological process

carotene metabolism | carotenoid metabolism | cell communication | ergosterol metabolism | ethylene metabolism | flower development | mevalonate pathway | photoprotection | ripening | xanthophyll metabolism

What is known about Lycopene?

Term details

Preferred name	lycopene
Alternative names	(6e,8e,10e,12e,14e,16e,18e,20e,22e,24e,26e)-2,6,8,10,12,14,16,18,20,22,24,26,30-tridecaene all-trans-lycopene psi,psi-carotene
Categories	metabolite
Nr of abstracts	2297

Actions:

[Retrieve abstracts in which term occurs](#)
[Add term to set](#)
[Search for indirect relations with term](#)

Below are the most significant co-occurrences of 'lycopene' with each category. Click on a category to view details.

[R scaled score >= 30](#)
[Abstract count >= 3](#)
[Order by](#)

disease

[Cancer of prostate](#) | [Eye diseases](#) | [Gastrointestinal neoplasm](#)
[Infertility, male](#) | [Maculopathy, age-related](#) | [Myoma](#)
[Vitamin a deficiency](#) | [cardiovascular disease](#) | [colon cancer](#) | [mouth neoplasm](#)

metabolite

[7,7',8,8',11,11',12,12'-octahydro-psi,psi-carotene](#) | [alpha-carotene](#) | [beta-zeacarotene](#) | [cyclic carotene](#)
[epsilon,epsilon-carotene](#) | [gamma-carotene](#) | [lutein](#)
[neurosporene](#) | [zeaxanthin](#) | [zeta-carotene](#)

pathway

[apoptosis](#) | [cell communication](#) | [cell cycle](#) | [dna repair](#) | [g1 phase](#)
[immune system](#) | [lipid metabolism](#) | [m phase](#) | [mevalonate pathway](#) | [s phase](#)

Abstracts with co-occurrences of 2 terms

Term 1 : [lycopene](#), [psi,psi-carotene](#), [all-trans-lycopene](#)

[\(6e,8e,10e,12e,14e,16e,18e,20e,22e,24e,26e\)-2,6,10,14,19,23,27,31-octamethyldotriacont-2,6,8,10,12,14,16,18,20,22,24,26,30-tridecaene](#)

Term 2 : [Maculopathy, age-related](#), [macular degeneration](#)

Terms co-occur together in 28 abstracts with an R scaled score of 42.

Page 1 of 3 | Most occur. first | Displaying 1 - 10 of 28

[Lycopene and lutein inhibit proliferation in rat prostate carcinoma cells.](#)

Richard S Gunasekera;Kiran Sewgobind;Smruti Desai;Larry Dunn;Homer S Black;Wallace L McKeehan;Bhimanagouda Patil;
 Nutr Cancer. 2007-07-20;58(2):171-7.
 PMID: 17640163

[Related citations](#)

Consumption of [lycopene](#), a carotenoid without provitamin A activity, has been associated with a lower risk of prostate and breast cancer. Lutein is another carotenoid that may be associated with a reduced risk of age-related [macular degeneration](#), the leading cause of blindness in adults 65 years of age and older. Bioactive compounds such as [lycopene](#) and lutein, derived from natural plant sources, have been shown to act at low substrate levels through the action of intrinsic cytokines and growth factors and their receptors within tissues, particularly those of the fibroblast growth factor and transforming growth factor beta families. The effects of grapefruit-derived and commercial [lycopene](#) and lutein preparations on androgen independent cultured malignant type II tumor cells [Dunning R3327AT3 or AT3 cells (androgen-responsive, slow-growing tumor cells with well developed epithelium and stroma)] were compared to their benign parent type I tumor epithelial cells (DTE). Results demonstrated that both [lycopene](#), in an alpha -cyclodextrin water soluble carrier, and lutein inhibited malignant AT3 cells in a concentration and time-dependent manner. No such effect was observed when benign DTE cells were examined, demonstrating selective inhibition of extremely malignant AT3 prostate cancer cells relative to their benign parent. Lutein demonstrated a similar but slightly diminished response as [lycopene](#). When cells were treated with cocktails of [lycopene](#) and lutein, no synergistic or additive effect occurred. These studies are consistent with epidemiological studies that show inverse relationships of these carotenoids with prostate cancer.

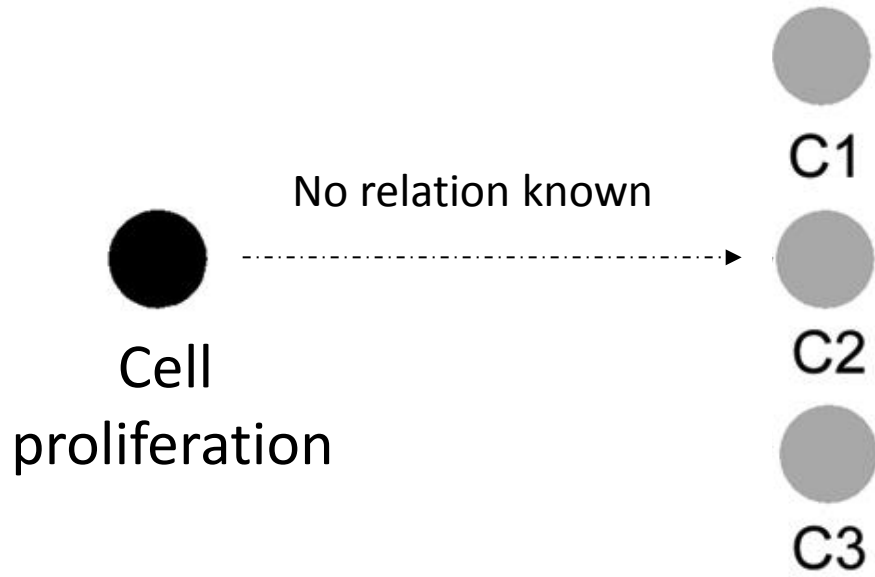
[Chemistry, distribution, and metabolism of tomato carotenoids and their impact on human health.](#)

Frederick Khachik;Lorena Carvalho;Paul S Bernstein;Garth J Muir;Da-You Zhao;Nikita B Katz;
 Exp Biol Med (Maywood). 2002-11-08;227(10):845-51.
 PMID: 12424324

[Related citations](#)

Recent epidemiological studies have suggested that the consumption of tomatoes and tomato-based food products reduce the risk of prostate cancer in humans. This protective effect has been attributed to carotenoids, which are one of the major classes of phytochemicals in this fruit. The most abundant carotenoid in tomato is [lycopene](#), followed by phytoene, phytofluene, zeta-carotene, gamma-carotene, beta-carotene, neurosporene, and lutein. The distribution of [lycopene](#) and related carotenoids in tomatoes and tomato-based food products has been determined by extraction and high-performance liquid chromatography-UV/visible photodiode array detection. Detailed qualitative and quantitative analysis of human serum, milk, and organs, particularly prostate, have revealed the presence of all the aforementioned carotenoids in biologically significant concentrations. Two oxidative metabolites of [lycopene](#), 2,6-cyclolycopene-

Discovery of new knowledge



Which existing drugs can be used to inhibit or activate cell proliferation?

Existing drugs

Swanson's
ABC principle
Fish oil cures Raynaud's disease

OPEN ACCESS Freely available online

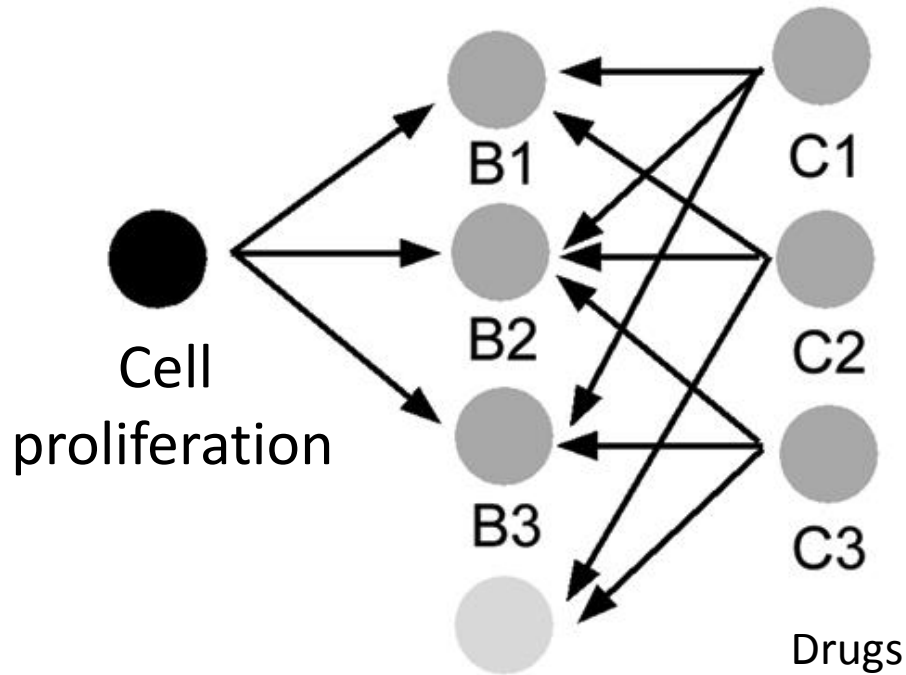
PLoS COMPUTATIONAL BIOLOGY

Literature Mining for the Discovery of Hidden Connections between Drugs, Genes and Diseases

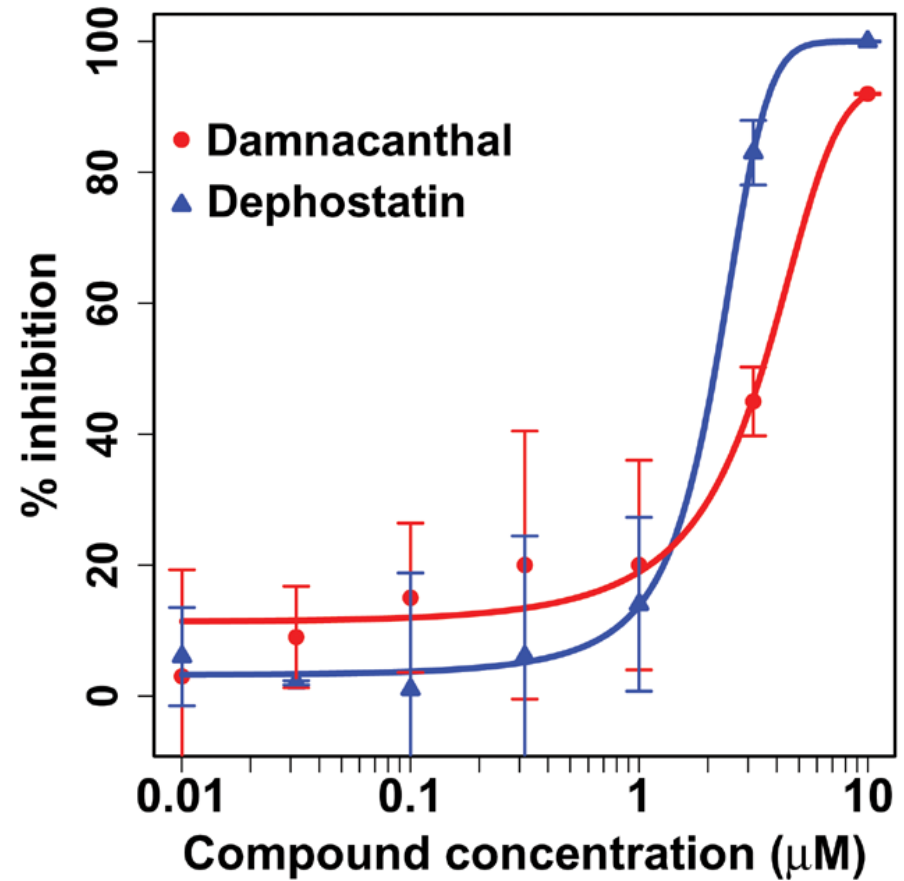
Raoul Frijters¹, Marianne van Vugt², Ruben Smeets², René van Schaik³, Jacob de Vlieg^{1,3}, Wynand Alkema^{3*}

1 Computational Drug Discovery (CDD), Nijmegen Centre for Molecular Life Sciences (NCMLS), Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands, 2 Department of Immune Therapeutics, Schering-Plough, Oss, The Netherlands, 3 Department of Molecular Design & Informatics, Schering-Plough, Oss, The Netherlands

Discovery of new knowledge



Damnacanthal
Dephostatin



Experimental validation

Linking to transcriptomics

Glucocorticoid regulated genes in the liver

Gene name	Symbol	Fc WT	Fc GRdim*
<i>Insulin-like growth factor binding protein 1</i>	<i>Igfbp1</i>	9.7	3.6
<i>Tyrosine aminotransferase</i>	<i>Tat</i>	4.1	1.6
<i>6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 3</i>	<i>Pfkfb3</i>	3.9	2.3
<i>Peroxisome proliferative activated receptor, gamma, coactivator 1 alpha</i>	<i>Ppargc1a</i>	3.2	3.2
<i>CCAAT/enhancer binding protein (C/EBP), beta</i>	<i>Cebpb</i>	3.1	1.6
<i>Protein-tyrosine sulfotransferase 2</i>	<i>Tpst2</i>	2.7	1.5
<i>Solute carrier family 2 (facilitated glucose transporter), member 1</i>	<i>Slc2a1</i>	2.6	1.9
<i>Forkhead box O1</i>	<i>Foxo1</i>	2.2	1.7
<i>Serine dehydratase</i>	<i>Sds</i>	2.1	1.5
<i>Aminoadipate-semialdehyde synthase</i>	<i>Aass</i>	2.1	1.3
<i>Phosphoenolpyruvate carboxykinase 1, cytosolic</i>	<i>Pck1</i>	1.9	1.3
<i>Tryptophan 2,3-dioxygenase</i>	<i>Tdo2</i>	1.5	1.2
<i>Aconitase 2, mitochondrial</i>	<i>Aco2</i>	1.4	1.3
<i>Transketolase</i>	<i>Tkt</i>	-1.2	-1.1
<i>Protein kinase, AMP-activated, beta 1 non-catalytic subunit</i>	<i>Prkab1</i>	-1.5	1
<i>Adiponectin receptor 2</i>	<i>Adipor2</i>	-1.6	-1.3
<i>Purinergic receptor P2Y, G-protein coupled, 5</i>	<i>P2ry5</i>	-1.6	-1.3
<i>Pyruvate kinase liver and red blood cell</i>	<i>Pklr</i>	-2.1	-1.4
<i>CCAAT/enhancer binding protein (C/EBP), alpha</i>	<i>Cebpa</i>	-2.3	-1.2
<i>Sterol regulatory element binding factor 1</i>	<i>Srebf1</i>	-3	-1.6
<i>Insulin receptor substrate 1</i>	<i>Irs1</i>	-4	-1.5

Linking to transcriptomics

Table 1: CoPub keyword enrichment analysis on prednisolone-regulated genes.

Biological process	# of associated genes
Metabolism	
Glucose metabolism/transport	22
Gluconeogenesis	21
Lipid metabolism/glycosylation	21
Glycolysis	14
Carbohydrate metabolism/transport	14
Amino acid metabolism/transport	13
Cell cycle and apoptosis	
Apoptosis	77
Cell proliferation	72
Cell cycle	65
Cell growth and-or maintenance, cell growth	63
Homeostasis	58
Cell differentiation	49
Cell cycle arrest	37
Immune/Inflammatory response	
Inflammatory response	31
Acute-phase response	20

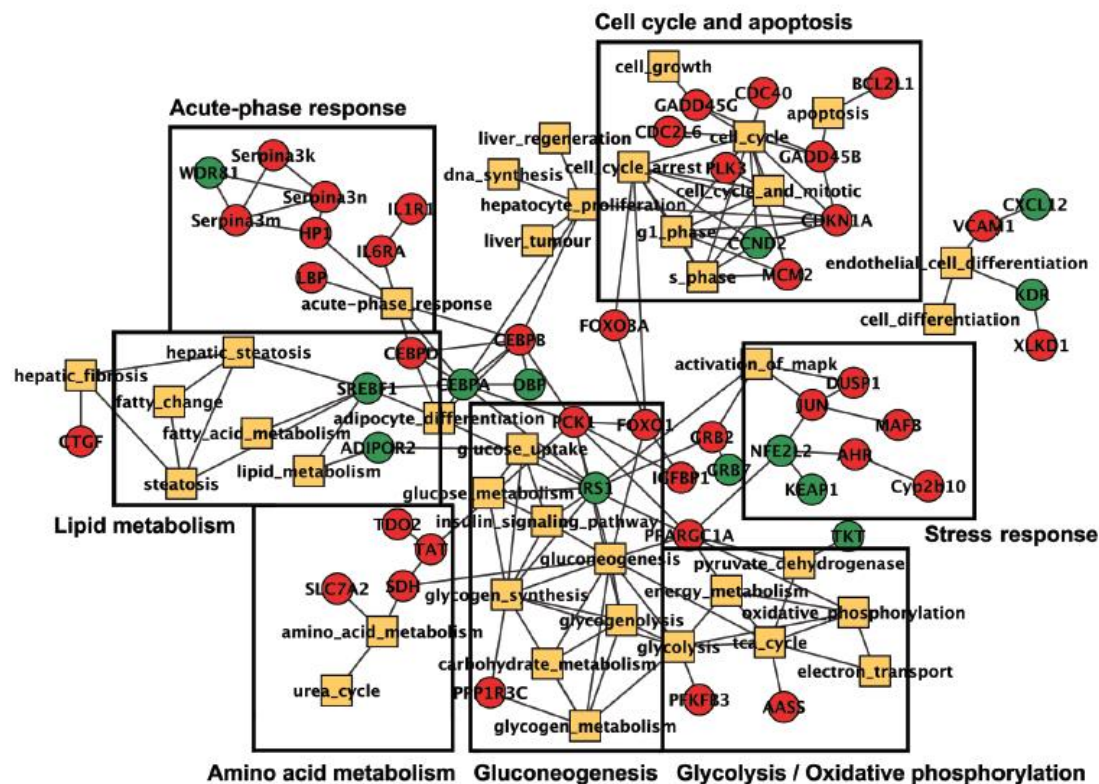
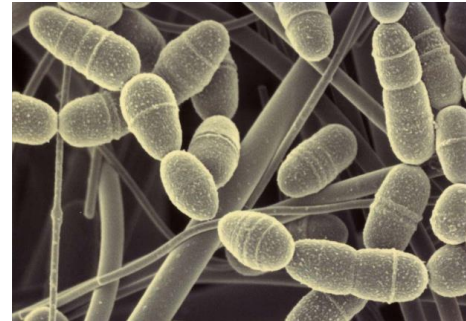
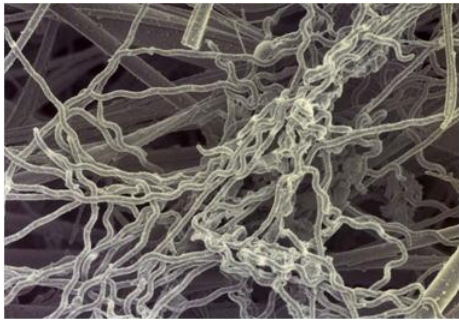
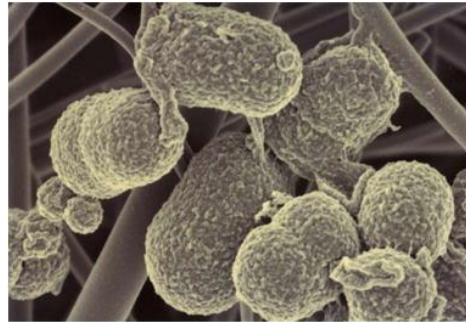
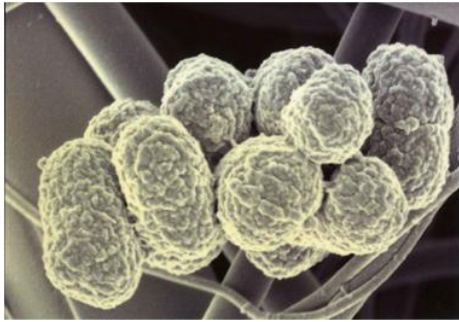


Figure 3 Literature-based network of glucocorticoid-induced effects. A network representation of the enrichment results was generated, in

You're mostly microbe

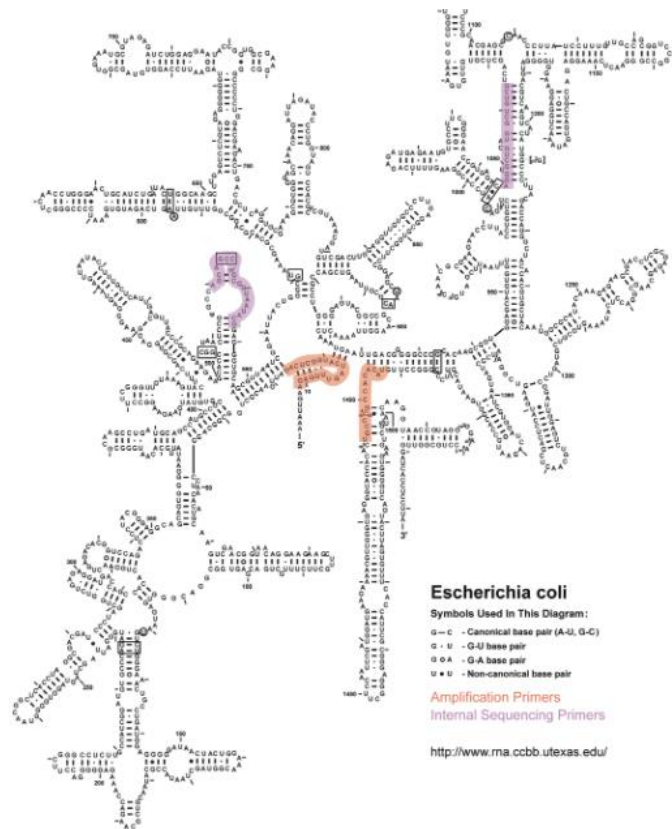


- 10^{18} cells, 10 times more than a human body
- In total more than 100 fold the gene content of a human
- Gut, skin, nasal-oral cavity
- Linked to health and disease

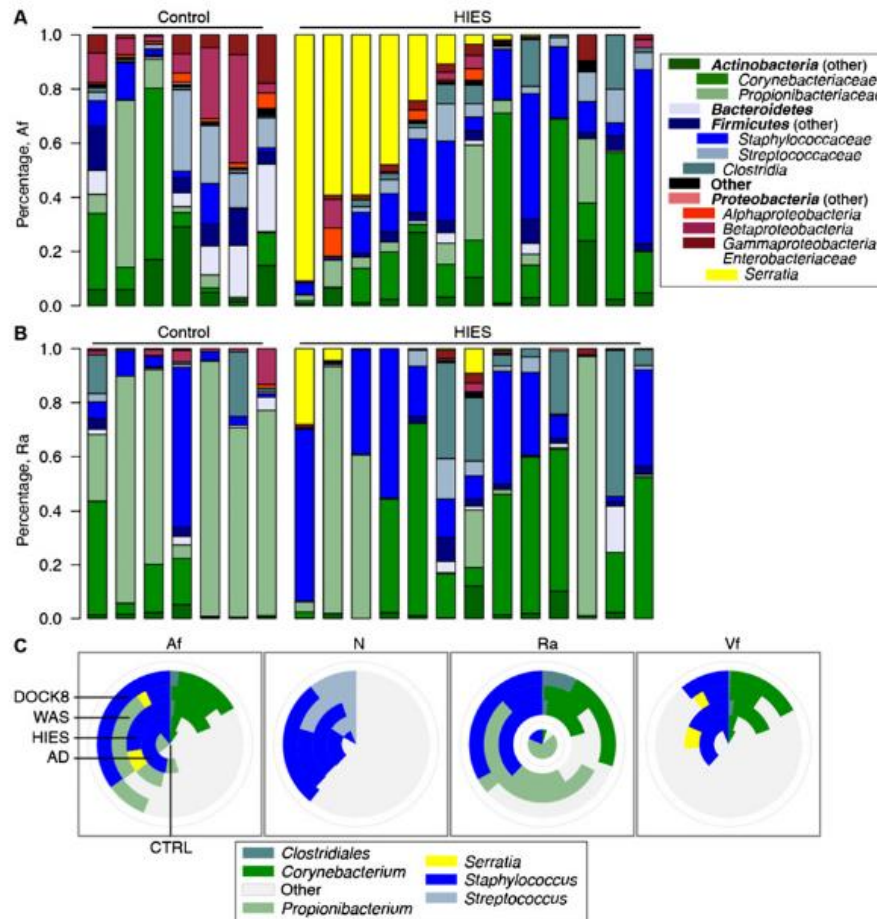
Metagenomics

“the use of genomics techniques to study communities of microbial organisms directly in their natural environments” (Chen and Pachter, 2005)

Profiling the skin microbiome : Who is there?



Secondary Structure: 16S small subunit ribosomal RNA



The altered landscape of the human skin microbiome in patients with primary immunodeficiencies

Julia Oh,¹ Alexandra F. Freeman,² NISC Comparative Sequencing Program,³ Morgan Park,³ Robert Sokolic,¹ Fabio Candotti,¹ Steven M. Holland,² Julia A. Segre,^{1,5,6} and Heidi H. Kong^{4,5,6}

¹Genetics and Molecular Biology Branch, National Human Genome Research Institute, NIH, Bethesda, Maryland 20892, USA; ²Laboratory of Clinical Infectious Diseases, National Institute of Allergy and Infectious Diseases, NIH, Bethesda, Maryland 20892, USA; ³NIH Intramural Sequencing Center, National Human Genome Research Institute, Bethesda, Maryland 20892, USA; ⁴Dermatology Branch, Center for Cancer Research, National Cancer Institute, NIH, Bethesda, Maryland 20892, USA

Profiling the skin microbiome : What do they do?

Table 2. Select abundant taxonomies differentially abundant between controls and STAT3-HIES individuals

Classification									
Site	Phylum/Order/Genus/Species			CTRL (mean ± SE)	STAT3-HIES (mean ± SE)	Adjusted P-value			
Af	Actinobacteria	Actinomycetales	<i>Corynebacterium</i>	9.1 ± 1.3%	17.2 ± 2.6%	0.04678			
			<i>Propionibacterium</i>	24.2 ± 3.0%	10.9 ± 2.7%	0.10688			
	Bacteroidetes	Flavobacteriales	<i>Porphyromonas</i>	0.9 ± 0.1%	0.3 ± 0.1%	0.04827			
			<i>Cloacibacterium</i>	3.0 ± 0.7%	0.0 ± 0.0%	0.04678			
	Firmicutes	Bacilliales	<i>Staphylococcus</i>	11.3 ± 1.7%	27.7 ± 4.0%	0.03092			
			<i>S. aureus</i>	0.4 ± 0.2%	10.3 ± 3.4%	0.02281			
			<i>S. epidermidis</i>	5.4 ± 1.4%	13.2 ± 2.7%	0.04893			
			<i>S. haemolyticus</i>	0.2 ± 0.0%	1.4 ± 0.4%	0.00050			
			Clostridiales	<i>Anaerococcus</i>	0.6 ± 0.1%	1.8 ± 0.4%	0.04480		
				<i>Fingoldia</i>	0.2 ± 0.0%	1.6 ± 0.4%	0.00001		
		Proteobacteria	Betaproteobacteria	<i>Peptoniphilus</i>	0.4 ± 0.1%	0.8 ± 0.2%	0.08678		
				<i>Diaphrobacter</i>	2.1 ± 0.6%	0.0 ± 0.0%	0.04678		
			Gammaproteobacteria	<i>Serratia</i>	0.0 ± 0.0%	8.1 ± 3.3%	0.00218		
		N	Firmicutes	Bacilliales	<i>Staphylococcus</i>	10.7 ± 1.3%	32.3 ± 4.3%	0.00002	
<i>S. aureus</i>	3.2 ± 0.9%				5.0 ± 1.9%	0.74406			
<i>S. epidermidis</i>	6.0 ± 0.9%				24.5 ± 4.2%	0.00000			
<i>S. haemolyticus</i>	0.0 ± 0.0%				0.5 ± 0.2%	0.02412			
<i>Streptococcus</i>	6.2 ± 1.1%				5.3 ± 1.4%	0.98904			
Ra	Actinobacteria				Actinomycetales	<i>Corynebacterium</i>	4.8 ± 1.2%	19.0 ± 4.5%	0.00163
						<i>Propionibacterium</i>	67.5 ± 4.8%	20.8 ± 6.8%	0.00004
	Firmicutes	Bacilliales	<i>Staphylococcus</i>	18.9 ± 3.6%	37.8 ± 5.2%	0.00556			
			<i>S. aureus</i>	0.0 ± 0.0%	9.1 ± 3.9%	0.00824			
			<i>S. epidermidis</i>	13.8 ± 2.9%	24.6 ± 5.4%	0.04219			
			<i>S. haemolyticus</i>	0.0 ± 0.0%	2.1 ± 1.6%	0.07403			
		<i>Fingoldia</i>	0.3 ± 0.1%	2.0 ± 0.6%	0.00197				
Vf	Actinobacteria	Actinomycetales	<i>Corynebacterium</i>	7.9 ± 1.4%	17.6 ± 2.6%	0.01084			
			<i>Cloacibacterium</i>	2.1 ± 0.6%	0.0 ± 0.0%	0.10420			
	Bacteroidetes	Flavobacteriales	<i>Porphyromonas</i>	1.2 ± 0.2%	0.5 ± 0.1%	0.07706			
			<i>Staphylococcus</i>	6.3 ± 1.1%	18.3 ± 2.7%	0.02686			
	Firmicutes	Bacilliales	<i>S. aureus</i>	0.6 ± 0.3%	7.5 ± 2.4%	0.05623			
			<i>S. epidermidis</i>	1.6 ± 0.2%	7.1 ± 1.5%	0.00002			
			<i>S. haemolyticus</i>	0.2 ± 0.1%	1.4 ± 0.4%	0.00042			
			Clostridiales	<i>Anaerococcus</i>	0.5 ± 0.1%	2.0 ± 0.4%	0.00001		
				<i>Fingoldia</i>	0.2 ± 0.0%	2.2 ± 0.5%	0.00000		
			Proteobacteria	Gammaproteobacteria	<i>Peptoniphilus</i>	0.3 ± 0.1%	1.2 ± 0.3%	0.00084	
		<i>Serratia</i>			0.0 ± 0.0%	8.4 ± 3.3%	0.00473		

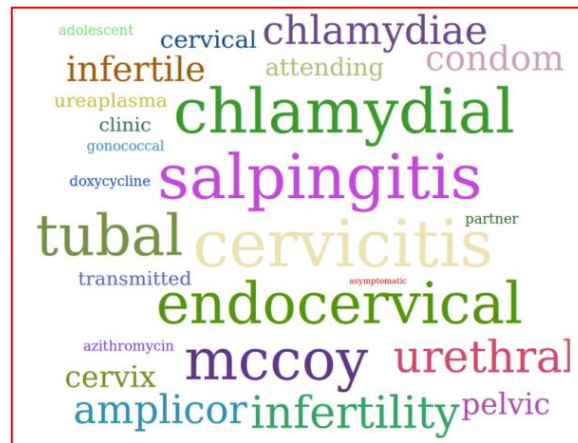
Microbes : What do they do?

Mapped entire NCBI taxonomy (~300000 species) against total MEDLINE

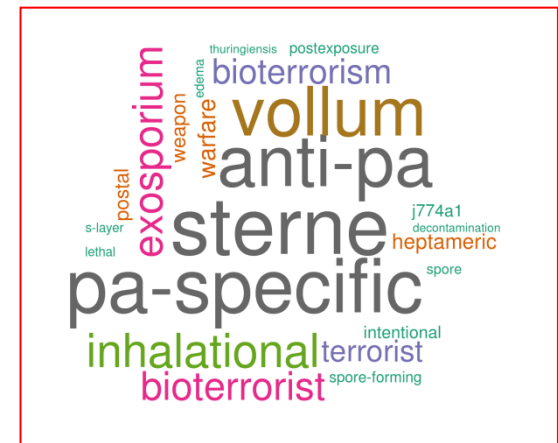
Unsupervised analysis : Count the words associated with your organism



Staphylococcus aureus
> 70000 scientific papers



Chlamydia
> 1200 scientific papers

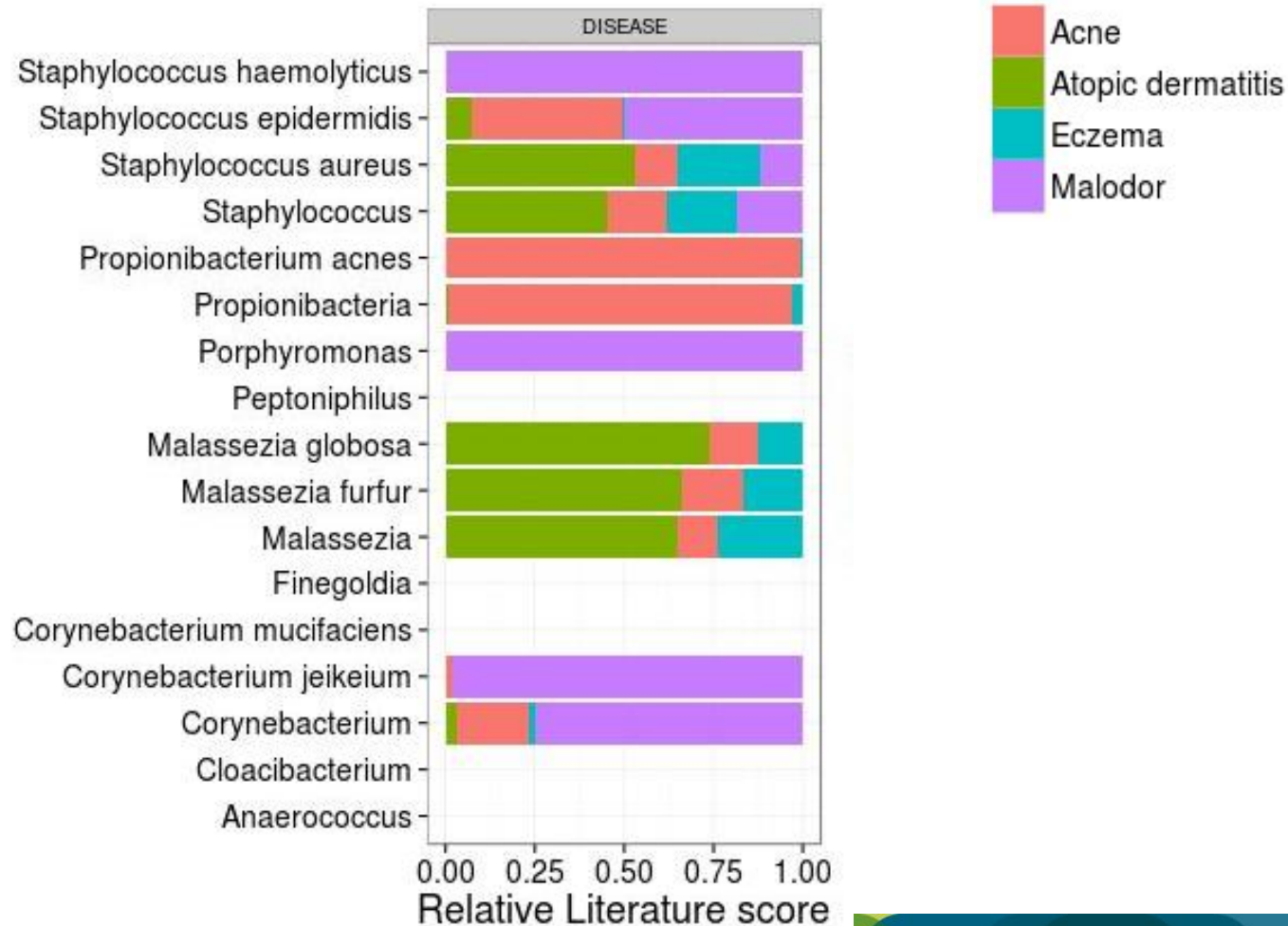


Bacillus anthracis
> 4000 scientific papers

Statistics, also when creating the word clouds

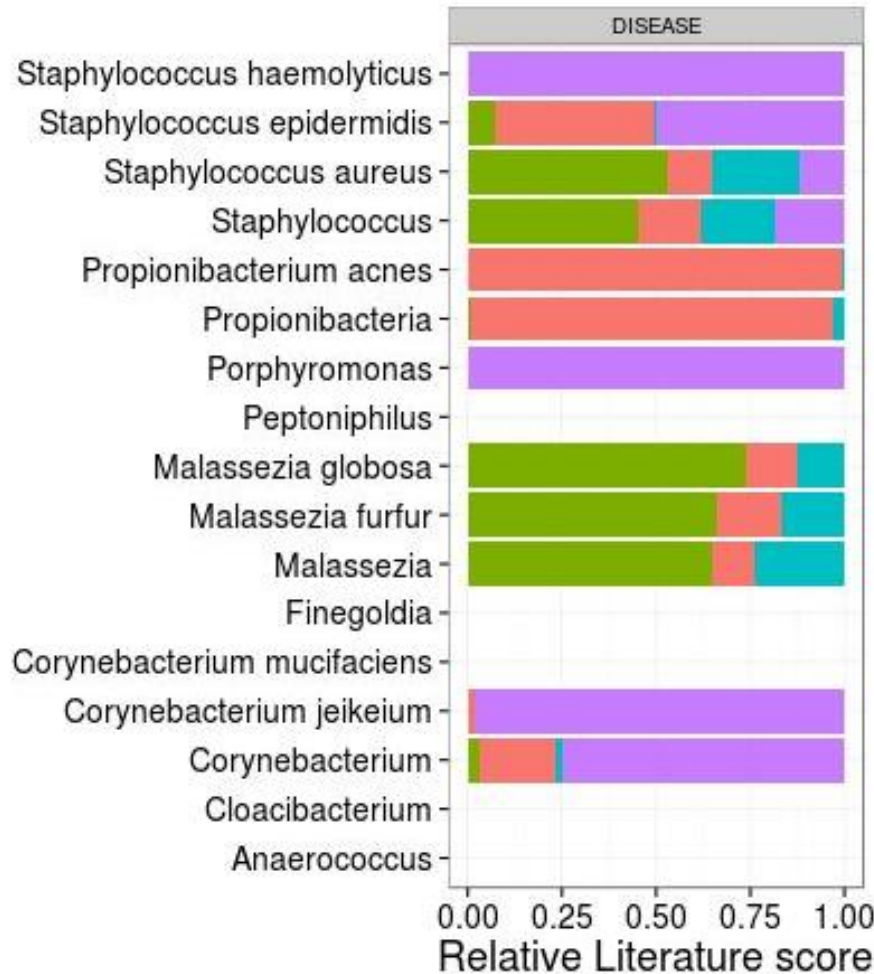
Linking the species to literature

Supervised analysis : Analyze known relations with diseases of interest



Linking the species to literature

Supervised analysis : Analyze known relations with diseases of interest



The in vivo effect of UVB radiation on skin bacteria in patients with atopic dermatitis.

Jekler, I M Bergbrant, J Faergemann, O Larkö. 1992-06-18. Acta Derm Venereol. 72. PMID: 1350138

Fourteen patients suffering from atopic dermatitis under treatment with UVB radiation were subjected to aerobic bacterial cultures in order to investigate whether this ultraviolet waveband has any in vivo germicidal effects, and, if so, whether there is a correlation with clinical improvement. Treatments were given 3 times a week for 8 weeks. Bacterial samples were collected before, midway and after the termination of therapy. On the latter two occasions, cultures were performed 10 min and 24 h post-UVB irradiation. The main bacteria found were Staphylococcus epidermidis and S. aureus. S. aureus carriage was found in 12 patients in lesional, dermatitic skin, and in 11 patients in clinically non-lesional skin. UVB radiation was found to have an antimicrobial effect primarily concerning S. aureus. Bacterial counts of this organism in lesional skin were decreased from a mean of $1.3 \times 10(3)$ to $1.2 \times 10(1)$ bacteria per cm^2 skin at the 8-week 30-min count (p less than 0.01) and $7.5 \times 10(1)$ at the 8-week 24-h count (p less than 0.05). The treatment yielded a statistically significant clinical improvement.

Antistaphylococcal IgE in patients with atopic dermatitis.

S Abramson, M V Dahl, G Walsh, M N Blumenthal, S D Douglas, P G Quie. 1982-10-21. J Am Acad Dermatol. 7. PMID: 7107990

Levels of IgE antibodies to Staphylococcus aureus and Staphylococcus epidermidis were determined in eleven patients with typical atopic dermatitis, with no history of furuncles or severe staphylococcal infection. Increased IgE binding to S. aureus but not to S. epidermidis was observed. Fifteen patients with hyperimmunoglobulinemia E-staphylococcal abscess syndrome had increased IgE binding not only to S. aureus but also to S. epidermidis. Other control groups of patients with elevated IgE levels or recurrent staphylococcal infection had normal IgE binding activity to both strains of staphylococci. Interaction of staphylococcal antigens from bacteria on skin with antistaphylococcal IgE antibodies on mast cells could induce mast cell release, evoke itch, and aggravate atopic dermatitis.

Comparative study of staphylococci from the skin of atopic dermatitis patients and from healthy subjects.

Higaki, M Morohashi, T Yamagishi, Y Hasegawa. 1999-06-17. Int J Dermatol. 38. PMID: 10321941

BACKGROUND: Bacterial infections occur frequently on the skin of atopic dermatitis (AD) patients. OBJECTIVE: This study was to evaluate the microbiology of the skin of AD patients for

Final remarks

- Technical issues in ‘bibliomics’ are largely solved
- Text mining is a versatile tool
 - Applicable to all kind of ~omics data.
 - Standardized workflows are needed!
 - Text mining on demand
- We need dedicated vocabularies that describe a set of concepts for a given topic, field of interest.
- Ontologies are good starting point but are not always practical (People do not use ontology terms in their papers)
- Algorithms for connecting concepts can be improved
 - Statistics
 - Network theory

Acknowledgements

