Florilège: une base de données intégrative de phénotypes microbiens grâce à l'ontologie Ontobiotope et aux traitements de text mining

E. Chaix, S. Dérozier, L. Deléger, H. Falentin, J.-B. Bohuon, M. Ba, R.Bossy, D. Sicard, V. Loux, C. Nédellec

Les écosystèmes microbiens

Quels microbes vivent dans quel environnement?

Techniques:

Méthode de culture

Génétique

Métagenomiques

Metatranscriptomique

Séquençage global (shotgun)

etc …

Helicobacter pylori
Mycobacterium avium
Escherichia coli ...

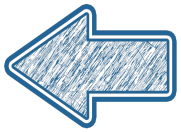
Legionella pneumophila
Yersinia pseudotuberculosis
Aeromonas hydrophila ...

Aspergillus flavus
Listeria seeligeri
Bacillus cereus ...

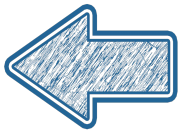
Quels sont les propriétés des habitats ? Des microbes?

Sources:

Documents textuels

- Publications scientifiques 
- Rapports
- Brevets
- Enregistrements médicaux

Bases de données

- Publiques : GenBank, SRA, Gold, CIRM INRA 
- Privées

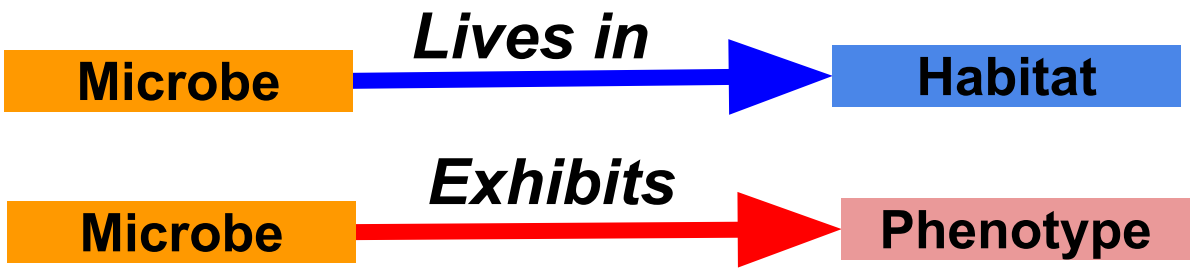
Données expérimentales

Modèle de connaissance :

Entités

Microbe	*Lactococcus fujiensis, Pseudomonas fluorescens, Listeria monocytogenes….*
Habitat	dairy cheese, beef, powdered milk ...
Phenotype	spore-forming, pathogenic,anaerobic ...

Relations



Croiser les informations

- **Taxon**
- **Habitat**
- **Phénotypes**

Difficulté : Formes très variables de textes dans les publications et BdD génomiques (GOLD, SRA, GenBank)

Croiser les informations

- Taxon
- Habitat
- Phénotypes

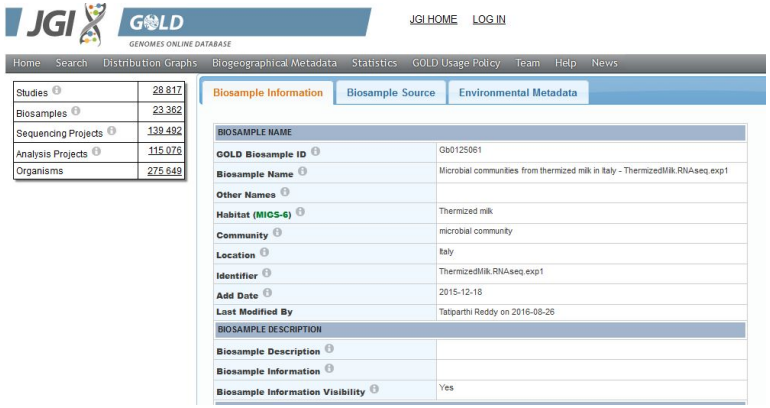
RESEARCH LETTER

Biogenic amine-forming microbial communities in cheese

Radka Burdychova & Tomas Komprda

Department of Food Technology, Mendel University of Agriculture and Forestry in Brno, Brno, Czech Republic

"Bacteria of the genera *Enterococcus* and *Lactobacillus* and coliform bacteria were isolated from Dutch-type semi-hard cheese"



ELSEVIER

International Journal of Food Microbiology 63 (2001) 91–98

INTERNATIONAL JOURNAL OF Food Microbiology

www.elsevier.nl/locate/ijfoodmicro

High incidence of *Listeria monocytogenes* in European red smear cheese

"Out of European red-smear cheese samples of various types [...] 1.2% of the samples were contaminated with *L. seeligeri*"

e.g.

- Artisanal cheeses from Tucuman
- Dairy cheese
- Caciocavallo cheese in Italy

Workflow de traitement

AlvisNLP

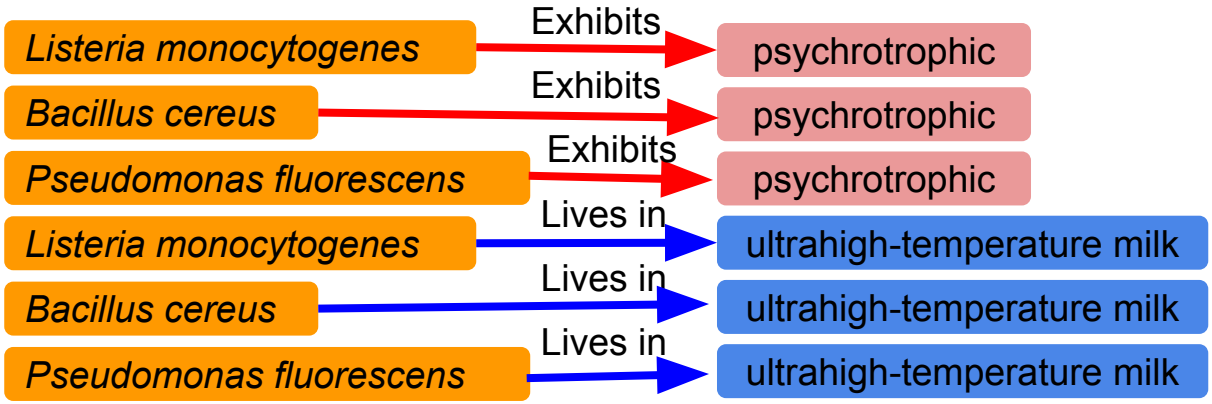
https://github.com/Bibliome/alvisnlp

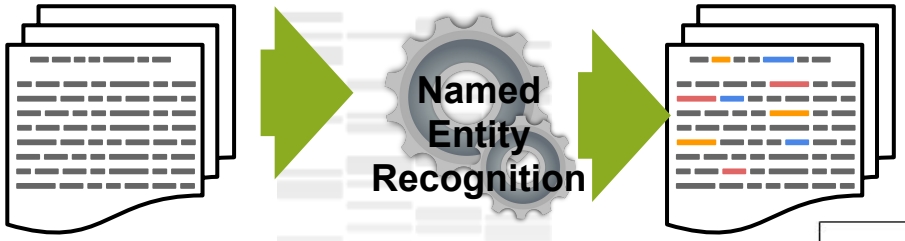
Workflow de traitement

AlvisNLP

https://github.com/Bibliome/alvisnlp

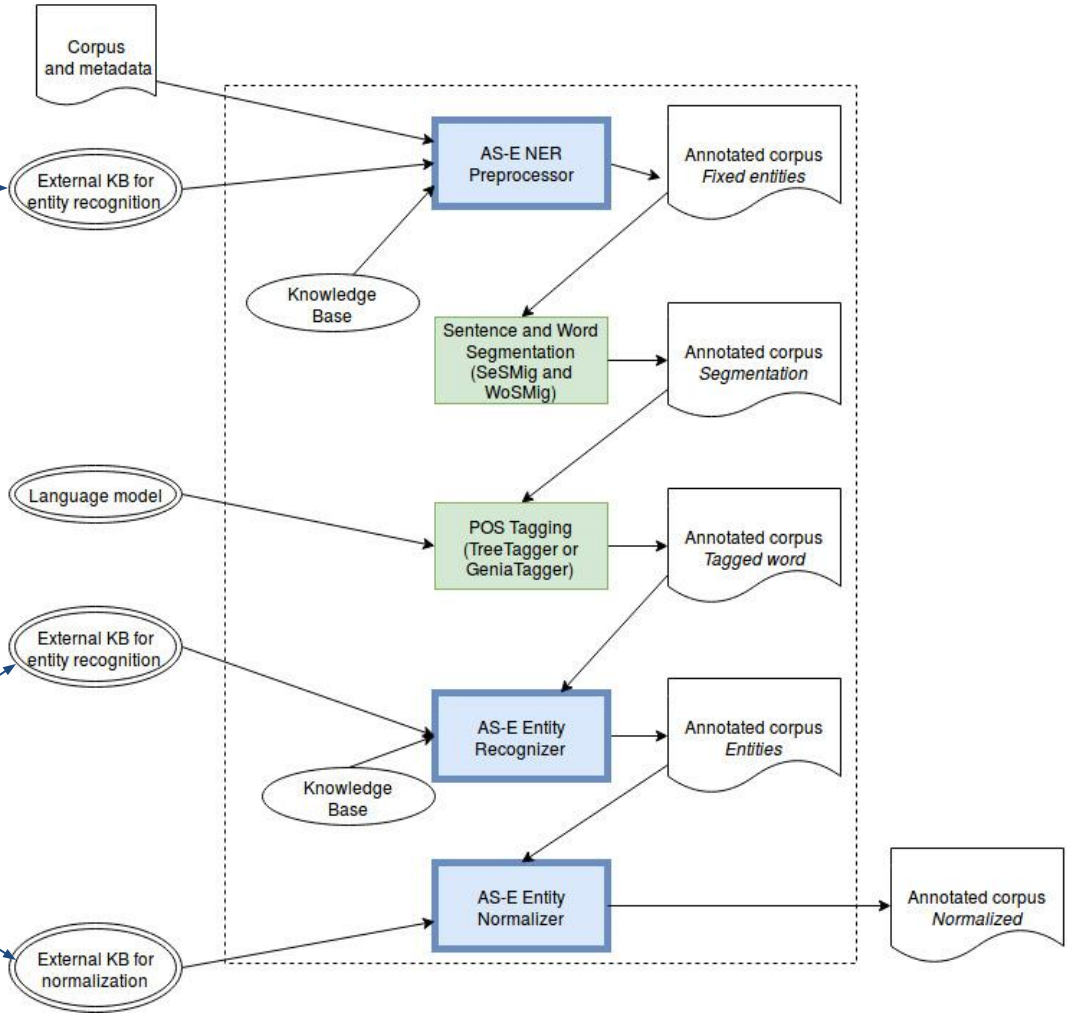
The effect of high hydrostatic pressure on the survival of the psychrotrophic organisms *Listeria monocytogenes*, *Bacillus cereus*, and *Pseudomonas fluorescens* was investigated in ultrahigh-temperature milk.





Lexiques, expressions régulières

Taxonomie et Ontologies



Ressources :

NCBI taxonomy

- Lactobacillaceae
  - Lactobacillus
    - Lactobacillus acetotolerans
      - Lactobacillus acetotolerans
    - Lactobacillus acidifarinae
      - Lactobacillus acidifarinae DSM 19394
    - Lactobacillus acidipiscis
      - Lactobacillus acidipiscis DSM 15836
      - Lactobacillus acidipiscis KCTC 13900
    - Lactobacillus acidophilus
      - Lactobacillus acidophilus 30SC
      - Lactobacillus acidophilus ATCC 4796
      - Lactobacillus acidophilus CFH
      - Lactobacillus acidophilus CIP 76.13
      - Lactobacillus acidophilus CIRM-BIA 442
      - Lactobacillus acidophilus CIRM-BIA 445
      - Lactobacillus acidophilus CRBIP 24179
      - Lactobacillus acidophilus DSM 20079
      - Lactobacillus acidophilus DSM 20242
      - Lactobacillus acidophilus DSM 9126
      - Lactobacillus acidophilus JV3179
      - Lactobacillus acidophilus La-14
      - Lactobacillus acidophilus NCFM

Microbe

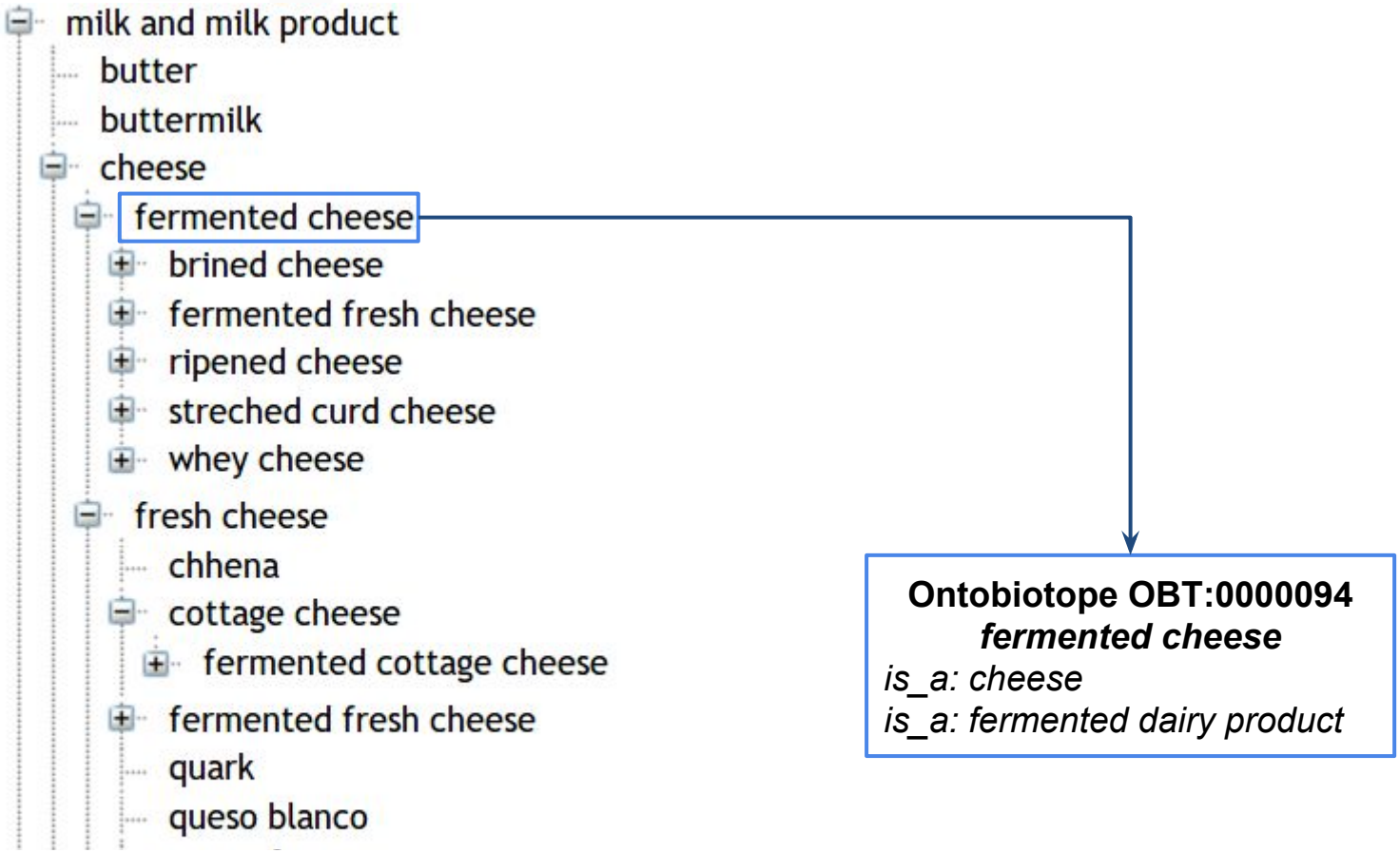
NCBI taxon 1579
Lactobacillus acidophilus
synonym: Thermobacterium intestinale, Bacillus acidophilus

Ressources :

http://agroportal.lirmm.fr/ontologies/ONTOBIOTOPE

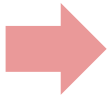
Ontobiotope Habitat

Habitat



Ressources :

Ontobiotope Phenotype



Phenotype

- microbial phenotype
  - phenotype wrt adhesion
  - phenotype wrt community behaviour
  - phenotype wrt growth
  - phenotype wrt metabolic activity
  - phenotype wrt microbial-host interaction
  - phenotype wrt morphology
  - phenotype wrt motility
    - motile
    - non motile
    - taxis phenotype
  - phenotype wrt ploidy
  - phenotype wrt stress
    - phenotype wrt chemical composition
      - acido resistant
      - acido sensitive
      - acido tolerant
      - alkali resistant

Ontobiotope OBT: 0000372
acido resistant
synonym: acidoresistant,acidresistant
is_a : phenotype wrt chemical composition
is_a: stress resistant

Extraction des termes

Groupe de mots d'intérêt : ici groupes nominaux et des adjectifs

Outil : YaTeA, qui utilise l'analyse syntaxique

The effect of high hydrostatic pressure on the survival of

the psychrotrophic organisms *Listeria monocytogene*s,

Bacillus cereus, and *Pseudomonas fluorescens*

was investigated in ultrahigh-temperature milk.

Extraction des termes

Groupe de mots d'intérêt : ici groupes nominaux et des adjectifs

Outil : YaTeA, qui utilise l'analyse syntaxique

The effect of high hydrostatic pressure on the survival of

DT	NN	IN	JJ	JJ	NN	IN	DT	NN	IN

the psychrotrophic organisms *Listeria monocytogenes*,

DT	JJ	NNS	NNP

Bacillus cereus, and *Pseudomonas fluorescens*

NNP	CC	NNP

was investigated in ultrahigh-temperature milk.

VBD	VBD	IN	JJ	NN

Termes candidats

- effect of high hydrostatic pressure
- high hydrostatic pressure
- hydrostatic pressure
- hydrostatic
- pressure
- survival of the psychrotrophic organisms
- psychrotrophic organisms
- psychrotrophic
- organisms
- Listeria monocytogenes
- Bacillus cereus
- Pseudomonas fluorescens
- ultrahigh-temperature milk
- ultrahigh-temperature
- milk

Catégorisation

Outil : ToMAP

Termes candidats

- effect of high hydrostatic pressure
- high hydrostatic pressure
- hydrostatic pressure
- hydrostatic
- pressure
- survival of the psychrotrophic organisms
- psychrotrophic organisms
- psychrotrophic
- organisms
- Listeria monocytogenes
- Bacillus cereus
- Pseudomonas fluorescens
- ultrahigh-temperature milk
- ultrahigh-temperature
- milk

Microbe

Stratégie de la catégorisation :
Comparaison des termes extraits et des labels des concepts de l'ontologie
- appariement exact
- appariement des têtes syntaxiques
- calcul de proximité sémantique (indice de Jaccard) entre :
- les termes extraits du texte
- les termes de l'ontologie

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|}.$$

ToMap : Golik, W.,et al. (2013). Improving term extraction with linguistic analysis in the biomedical domain. https://frama.link/Golik2011

Catégorisation

Outil : ToMAP

Extraits de Ontobiotope

Termes candidats

- effect of high hydrostatic pressure
- high hydrostatic pressure
- hydrostatic pressure
- hydrostatic
- pressure
- survival of the psychrotrophic organisms
- psychrotrophic organisms
- **psychrotrophic**
- organisms
- Listeria monocytogenes
- Bacillus cereus
- Pseudomonas fluorescens
- **ultrahigh-temperature milk**
- ultrahigh-temperature
- **milk**

Microbe

appariement exact

→ microbial phenotype
→ physiological phenotype
→ phentoype wrt habitat temperature
→ cryophile
→ mesophile
→ psychrophile
→ psychrotrophic
→ thermophile

→ microbial habitat
→ food
→ animal product and primary derivative thereof
→ milk and milk product
→ butter
→ cheese
→ ice cream
→ milk
→ yogurt

appariement des têtes syntaxiques

appariement exact

ToMap : Golik, W.,et al. (2013). Improving term extraction with linguistic analysis in the biomedical domain. https://frama.link/Golik2011

Choix dans la catégorisation

The effect of high hydrostatic pressure on the survival of the psychrotrophic organisms *Listeria monocytogene*s, *Bacillus cereus*, and *Pseudomonas fluorescens* was investigated in ultrahigh-temperature milk.

→ **Si deux termes imbriqués sont étiquetés par la même catégorie, on ne conserve que le plus grand**

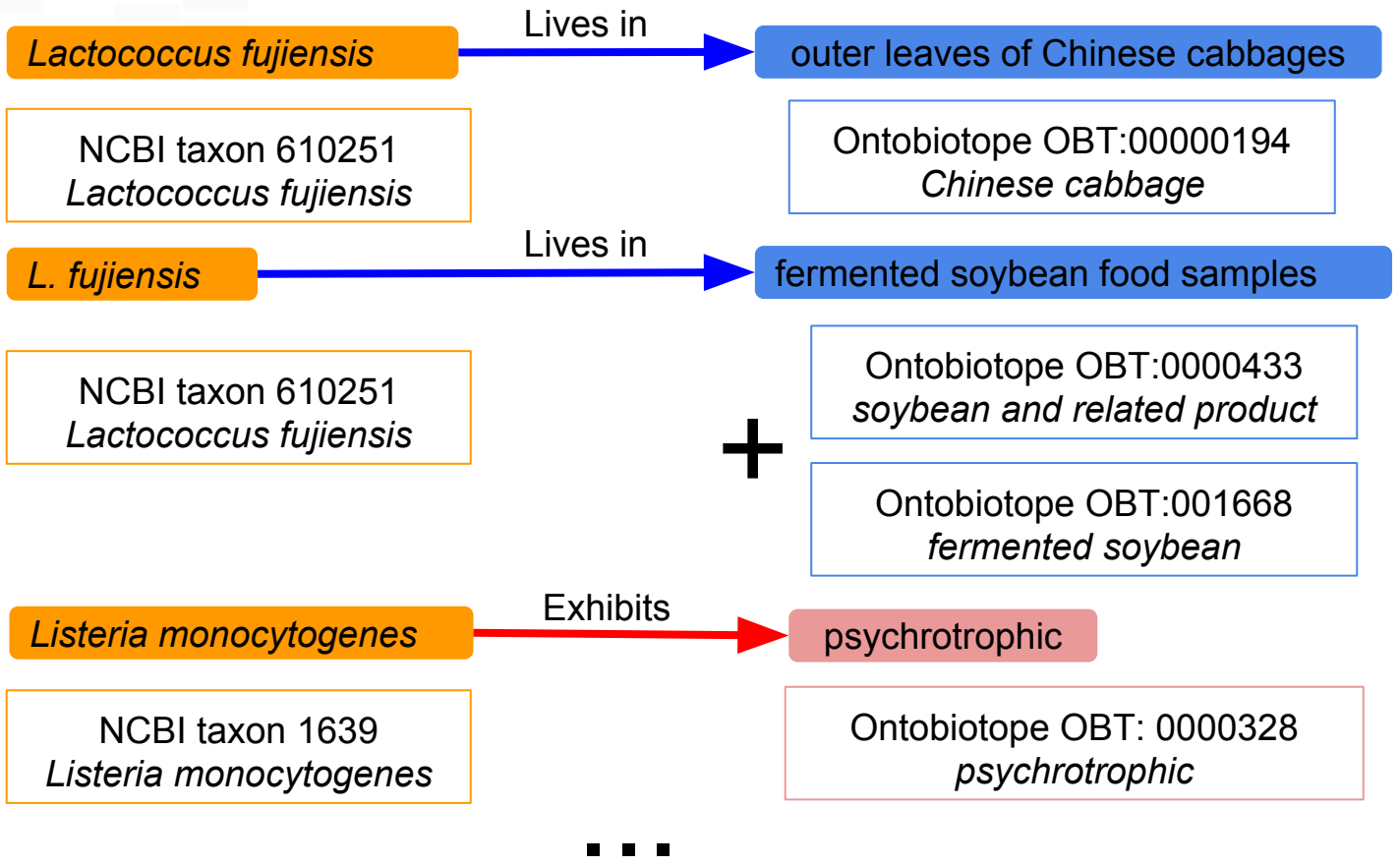
e.g. ultrahigh-temperature milk* et *milk* taggués par *milk

→ **Score de proximité : si plusieurs termes de l'ontologie possèdent la même tête, catégorisation par la classe la plus spécifique**

→ **Possibilité de catégoriser par plusieurs catégories**

e.g. ultrahigh-temperature milk* taggués par milk et *UHT food
traditional fermented cabbage* par *cabbage* et *fermented food

En sortie du traitement de TDM



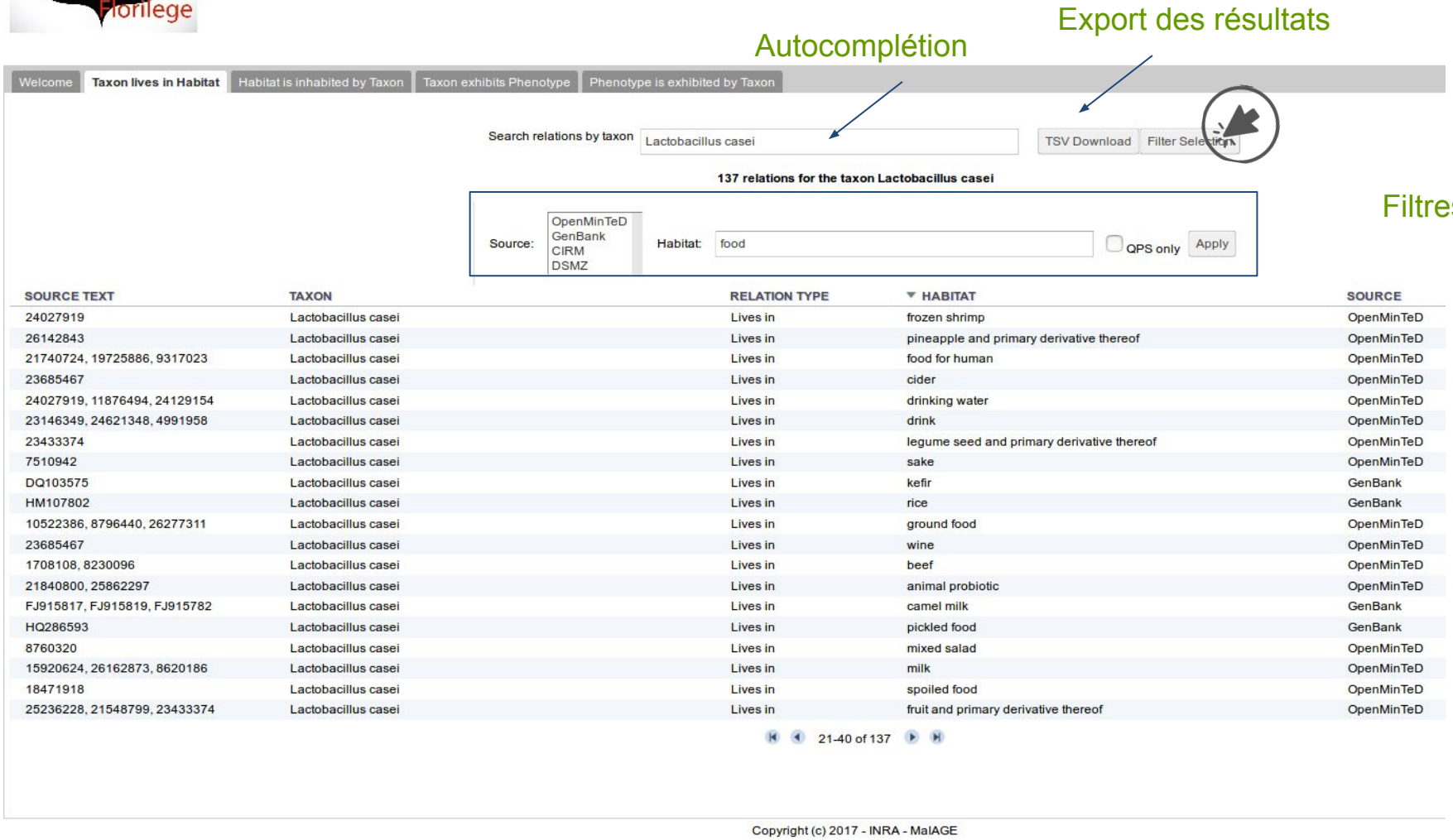
Florilège : application web

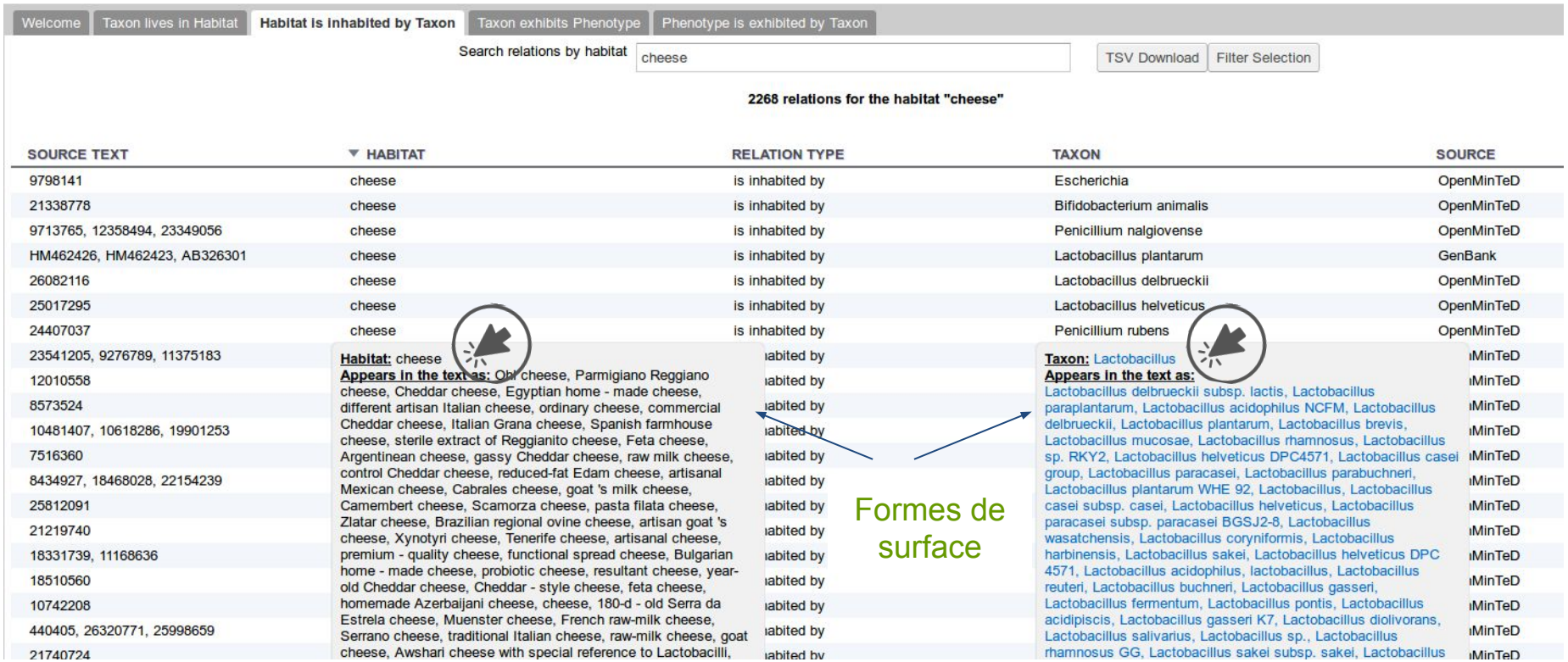
http://migale.jouy.inra.fr/Florilege

- Taxon ↔ Habitat : 820 000 relations
- Taxon ↔ Phenotype : 86 000 relations
- Lien direct avec les données publiques externes
  - PubMed (2.10^6 résumés), DSMZ (>60 000 notices), GenBank (~60 000 entrées)...
- Filtres pour sélectionner la source de données, le statut QPS …
- Export possible au format tabulé

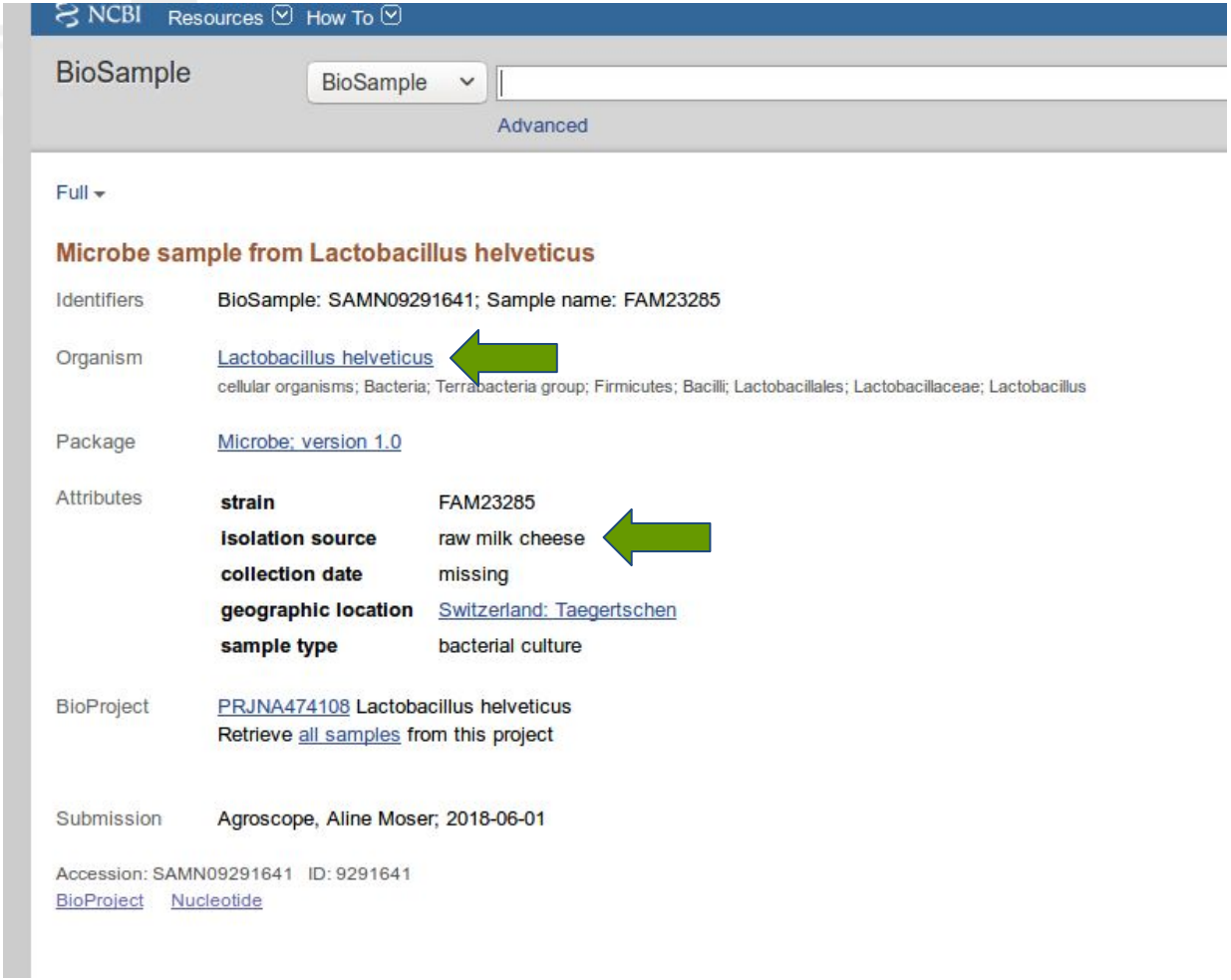
Florilège : application web

http://migale.jouy.inra.fr/Florilege

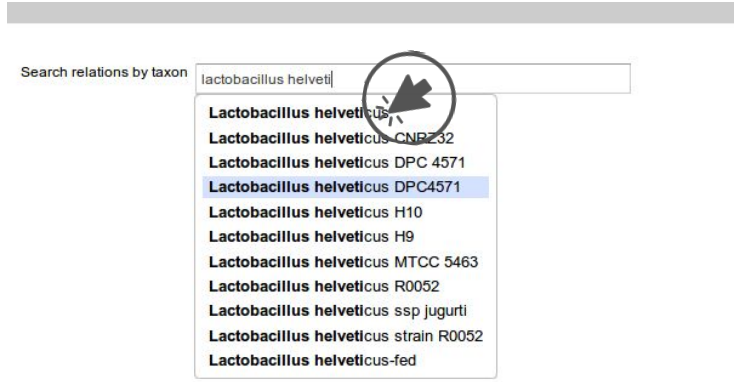

Export des résultats
Autocomplétion
Welcome
Taxon lives in Habitat
Habitat is inhabited by Taxon
Taxon exhibits Phenotype
Phenotype is exhibited by Taxon
Search relations by taxon
Lactobacillus casei
TSV Download
Filter Selection
137 relations for the taxon Lactobacillus casei
Filtres
Source:
OpenMinTeD
GenBank
CIRM
DSMZ
Habitat:
food
QPS only
Apply
SOURCE TEXT
TAXON
RELATION TYPE
HABITAT
SOURCE
24027919 Lactobacillus casei Lives in frozen shrimp OpenMinTeD
26142843 Lactobacillus casei Lives in pineapple and primary derivative thereof OpenMinTeD
21740724, 19725886, 9317023 Lactobacillus casei Lives in food for human OpenMinTeD
23685467 Lactobacillus casei Lives in cider OpenMinTeD
24027919, 11876494, 24129154 Lactobacillus casei Lives in drinking water OpenMinTeD
23146349, 24621348, 4991958 Lactobacillus casei Lives in drink OpenMinTeD
23433374 Lactobacillus casei Lives in legume seed and primary derivative thereof OpenMinTeD
7510942 Lactobacillus casei Lives in sake OpenMinTeD
DQ103575 Lactobacillus casei Lives in kefir GenBank
HM107802 Lactobacillus casei Lives in rice GenBank
10522386, 8796440, 26277311 Lactobacillus casei Lives in ground food OpenMinTeD
23685467 Lactobacillus casei Lives in wine OpenMinTeD
1708108, 8230096 Lactobacillus casei Lives in beef OpenMinTeD
21840800, 25862297 Lactobacillus casei Lives in animal probiotic OpenMinTeD
FJ915817, FJ915819, FJ915782 Lactobacillus casei Lives in camel milk GenBank
HQ286593 Lactobacillus casei Lives in pickled food GenBank
8760320 Lactobacillus casei Lives in mixed salad OpenMinTeD
15920624, 26162873, 8620186 Lactobacillus casei Lives in milk OpenMinTeD
18471918 Lactobacillus casei Lives in spoiled food OpenMinTeD
25236228, 21548799, 23433374 Lactobacillus casei Lives in fruit and primary derivative thereof OpenMinTeD
21-40 of 137
Copyright (c) 2017 - INRA - MaIAGE


Welcome
Taxon lives in Habitat
Habitat is inhabited by Taxon
Taxon exhibits Phenotype
Phenotype is exhibited by Taxon
Search relations by habitat
cheese
TSV Download
Filter Selection
2268 relations for the habitat "cheese"
SOURCE TEXT
HABITAT
RELATION TYPE
TAXON
SOURCE
9798141 cheese is inhabited by Escherichia OpenMinTeD
21338778 cheese is inhabited by Bifidobacterium animalis OpenMinTeD
9713765, 12358494, 23349056 cheese is inhabited by Penicillium nalgiovense OpenMinTeD
HM462426, HM462423, AB326301 cheese is inhabited by Lactobacillus plantarum GenBank
26082116 cheese is inhabited by Lactobacillus delbrueckii OpenMinTeD
25017295 cheese is inhabited by Lactobacillus helveticus OpenMinTeD
24407037 cheese is inhabited by Penicillium rubens OpenMinTeD
Habitat: cheese
Appears in the text as: Obi cheese, Parmigiano Reggiano cheese, Cheddar cheese, Egyptian home - made cheese, different artisan Italian cheese, ordinary cheese, commercial Cheddar cheese, Italian Grana cheese, Spanish farmhouse cheese, sterile extract of Reggianito cheese, Feta cheese, Argentinean cheese, gassy Cheddar cheese, raw milk cheese, control Cheddar cheese, reduced-fat Edam cheese, artisanal Mexican cheese, Cabrales cheese, goat 's milk cheese, Camembert cheese, Scamorza cheese, pasta filata cheese, Zlatar cheese, Brazilian regional ovine cheese, artisan goat 's cheese, Xynotyri cheese, Tenerife cheese, artisanal cheese, premium - quality cheese, functional spread cheese, Bulgarian home - made cheese, probiotic cheese, resultant cheese, year-old Cheddar cheese, Cheddar - style cheese, feta cheese, homemade Azerbaijani cheese, cheese, 180-d - old Serra da Estrela cheese, Muenster cheese, French raw-milk cheese, Serrano cheese, traditional Italian cheese, raw-milk cheese, goat cheese, Awshari cheese with special reference to Lactobacilli,
Formes de surface
Taxon: Lactobacillus
Appears in the text as:
Lactobacillus delbrueckii subsp. lactis, Lactobacillus paraplantarum, Lactobacillus acidophilus NCFM, Lactobacillus delbrueckii, Lactobacillus plantarum, Lactobacillus brevis, Lactobacillus mucosae, Lactobacillus rhamnosus, Lactobacillus sp. RKY2, Lactobacillus helveticus DPC4571, Lactobacillus casei group, Lactobacillus paracasei, Lactobacillus parabuchneri, Lactobacillus plantarum WHE 92, Lactobacillus, Lactobacillus casei subsp. casei, Lactobacillus helveticus, Lactobacillus paracasei subsp. paracasei BGSJ2-8, Lactobacillus wasatchensis, Lactobacillus coryniformis, Lactobacillus harbinensis, Lactobacillus sakei, Lactobacillus helveticus DPC 4571, Lactobacillus acidophilus, lactobacillus, Lactobacillus reuteri, Lactobacillus buchneri, Lactobacillus gasseri, Lactobacillus fermentum, Lactobacillus pontis, Lactobacillus acidipiscis, Lactobacillus gasseri K7, Lactobacillus diolivorans, Lactobacillus salivarius, Lactobacillus sp., Lactobacillus rhamnosus GG, Lactobacillus sakei subsp. sakei, Lactobacillus

Utilisation 1 : je trouve une bactérie dans un échantillon est-ce normal?



Lactobacillus helveticus



Welcome | **Taxon lives in Habitat** | Habitat is inhabited by Taxon

Search relations by taxon: Lactobacillus helveticus — TSV Download — Filter Selection

255 relations for the taxon Lactobacillus helveticus

DOCUMENT	TAXON	RELATION TYPE	HABITAT	SOURCE
PMID: 3924282	Lactobacillus helveticus	Lives in	urine	OpenMinTeD
PMID: 6374570	Lactobacillus helveticus	Lives in	gum tissue	OpenMinTeD
PMID: 10607224, 12004752, 15714491	Lactobacillus helveticus	Lives in	pregnant woman	OpenMinTeD
PMID: 25306273	Lactobacillus helveticus H9	Lives in	fermented milk	OpenMinTeD
PMID: 20308300, 20974015, 21983070	Lactobacillus helveticus R0052	Lives in	animal probiotic	OpenMinTeD
PMID: 24935567	Lactobacillus helveticus	Lives in	sugar	OpenMinTeD
PMID: 12755477, 15334032, 16246229	Lactobacillus helveticus	Lives in	rat	OpenMinTeD
PMID: 19050899	Lactobacillus helveticus	Lives in	embryo	OpenMinTeD
PMID: 24905205	Lactobacillus helveticus DPC 4571	Lives in	gut	OpenMinTeD
PMID: 3537138	Lactobacillus helveticus	Lives in	horseradish	OpenMinTeD
PMID: 11097908, 11895556, 16246229	Lactobacillus helveticus	Lives in	combustible liquid	OpenMinTeD
PMID: 25151888	Lactobacillus helveticus	Lives in	fermented food	OpenMinTeD
PMID: 11860905, 16323084, 25965250	Lactobacillus helveticus	Lives in	patient with carcinoma	OpenMinTeD
PMID: 1462050, 16148174, 16419579	Lactobacillus helveticus	Lives in	animal	OpenMinTeD
PMID: 17594806, 19804649	Lactobacillus helveticus	Lives in	sandwich	OpenMinTeD
PMID: 11928956, 17184869, 17336831	Lactobacillus helveticus	Lives in	mucosal tissue	OpenMinTeD
PMID: 16987432, 24905205	Lactobacillus helveticus DPC 4571	Lives in	cheese	OpenMinTeD
PMID: -	Lactobacillus helveticus	Lives in	milk and milk product	GenBank
PMID: 9678685	Lactobacillus helveticus	Lives in	muscle	OpenMinTeD
PMID: 16297479, 25892687	Lactobacillus helveticus	Lives in	cow milk	OpenMinTeD

21-40 of 255

NCBI Resources How To

PubMed.gov
US National Library of Medicine
National Institutes of Health

PubMed

Advanced

Format: Abstract

Send to

J Dairy Sci. 2014 Dec;97(12):7413-25. doi: 10.3168/jds.2014-8520. Epub 2014 Oct 11.

Proteome analysis of Lactobacillus helveticus H9 during growth in skim milk.

Chen YF[1], Zhao WJ[2], Wu RN[3], Sun ZH[1], Zhang WY[4], Wang JC[4], Bilige M[4], Zhang HP[5].

Author information

Abstract

Lactobacillus helveticus H9 was isolated from traditionally fermented yak milk in Tibet (China) with the ability to produce the antihypertensive peptides Val-Pro-Pro (VPP) and Ile-Pro-Pro (IPP) during milk fermentation. To understand the changes in the protein expression of L. helveticus H9, proteome analysis was performed at 3 different growth stages, lag phase (pH 6.1), log phase (pH 5.1), and stationary phase (pH 4.5) using 2-dimensional electrophoresis (2-DE). Further analysis showed that 257 differential protein spots were found and 214 protein spots were identified using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF/MS). The cellular functions of the differentially expressed proteins were complex. Interestingly, the proteolytic system-related proteins aminopeptidase N (PepN), aminopeptidase E (PepE), endopeptidase O2 (PepO2), and oligopeptide transport system permease protein (OppC) were observed only on the maps of pH 5.1 and pH 4.5, which was consistent with the presence of angiotensin I-converting enzyme (ACE)-inhibitory peptides VPP and IPP during these 2 growth stages (log phase and stationary phase). These results, combined with a previous study of gene expression of the proteolytic system, led us to conclude that the Opp transport system, pepE, and pepO2 are likely related to the production of ACE-inhibitory peptides.

KEYWORDS: Lactobacillus helveticus H9; proteolytic system; proteomics

PMID: 25306273 DOI: 10.3168/jds.2014-8520

[Indexed for MEDLINE]

Utilisation 2 : aide à l'innovation alimentaire

je veux promouvoir le sel de Guérande à travers un nouveau produit fermenté

→ je cherche une bactérie QPS halotolérante :

Welcome | Taxon lives in Habitat | Habitat is inhabited by Taxon | Taxon exhibits Phenotype | **Phenotype is exhibited by Taxon**

Search relations by phenotype: halotolerant — TSV Download — Filter Selection

9 relations for the phenotype "halotolerant"

Source: OpenMinTeD, GenBank, CIRM, DSMZ — Taxon: — QPS only — Apply

SOURCE TEXT	PHENOTYPE	RELATION TYPE	TAXON	SOURCE
9327565	halotolerant	is exhibited by	Saccharomyces cerevisiae	OpenMinTeD
25039289	halotolerant	is exhibited by	Lactobacillus plantarum	OpenMinTeD
17897213, 25542205	halotolerant	is exhibited by	Bacillus pumilus	OpenMinTeD
15849794, 16467467, 17072537	halotolerant	is exhibited by	Debaryomyces hansenii	OpenMinTeD
18068256	halotolerant	is exhibited by	Lactococcus lactis	OpenMinTeD
12486459, 12557391, 21890005	halotolerant	is exhibited by	Bacillus subtilis	OpenMinTeD
18068256	halotolerant	is exhibited by	Leuconostoc lactis	OpenMinTeD
21664643	halotolerant	is exhibited by	Bacillus megaterium	OpenMinTeD
25561404, 16488097, 7646007	halotolerant	is exhibited by	Bacillus licheniformis	OpenMinTeD

1-9 of 9

Où vit-elle ?

Welcome | **Taxon lives in Habitat** | Habitat is inhabited by Taxon | Taxon exhibits Phenotype | Phenotype is exhibited by Taxon

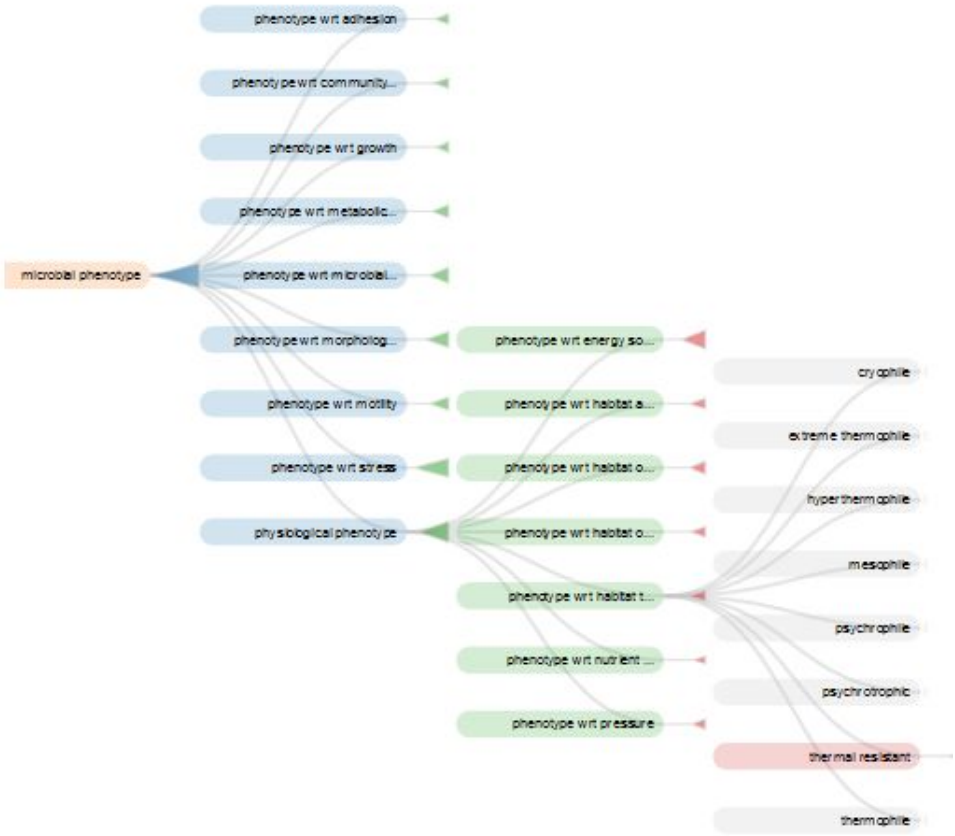
EF420763	Lactobacillus plantarum	Lives in	animal	GenBank
GU552552	Lactobacillus plantarum	Lives in	baby corn and related product	GenBank
AB526458, AB526455, AB526454	Lactobacillus plantarum	Lives in	bamboo shoot	GenBank
GQ289387	Lactobacillus plantarum	Lives in	beer	GenBank
GU138591, GU138574, GU138589	Lactobacillus plantarum	Lives in	bread	GenBank
AM157432	Lactobacillus plantarum	Lives in	breast milk	GenBank
AM157432	Lactobacillus plantarum	Lives in	breast	GenBank
6496	Lactobacillus plantarum	Lives in	brine	DSMZ
EF422373	Lactobacillus plantarum	Lives in	cabbage	GenBank
6493	Lactobacillus plantarum	Lives in	cabbage	DSMZ
6629	Lactobacillus plantarum	Lives in	cabbage	DSMZ
FJ751793	Lactobacillus plantarum	Lives in	calf	GenBank
FJ915735, FJ915814, FJ915815	Lactobacillus plantarum	Lives in	camel milk	GenBank
FJ538531, FJ538504, FJ538512	Lactobacillus plantarum	Lives in	cattle	GenBank
HM462426, HM462423, AB326301	Lactobacillus plantarum	Lives in	cheese	GenBank
HM218104, HM218301, FJ607272	Lactobacillus plantarum	Lives in	dairy industry	GenBank
FJ227310, FJ227315	Lactobacillus plantarum	Lives in	drink	GenBank

1-20 of 622

Fromage au lait de chamelle ou maïs fermenté?

Perspectives

- Mise à jour de Florilège intégrant les Phénotypes
- MàJ d'Ontobiotope sur Agroportal
  - Branche "Phénotype"
- Recherche et Navigation par Ontologie dans Florilège



Merci pour votre attention