

Text mining tools for extracting information about microbial biodiversity in food

Estelle Chaix, Louise Deleger, Robert Bossy, Claire Nédellec

▶ To cite this version:

Estelle Chaix, Louise Deleger, Robert Bossy, Claire Nédellec. Text mining tools for extracting information about microbial biodiversity in food. Microbial Spoilers in Food 2017, Association pour le Développement de la Recherche Appliquée aux Industries Agricoles et Alimentaires (ADRIA). FRA., Jun 2017, Quimper, France. pp.95. hal-01602552

HAL Id: hal-01602552 https://hal.science/hal-01602552

Submitted on 2 Jun2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Text mining tools for extracting information about microbial biodiversity in food

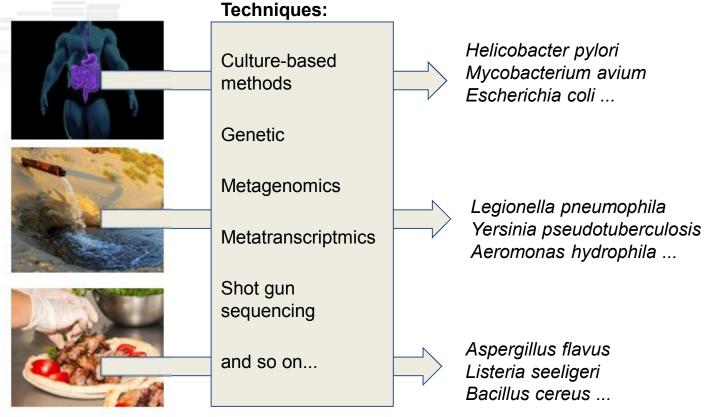
Bibliome Team MaIAGE, INRA, Université Paris-Saclay, 78350 Jouy-en-Josas, France

Estelle Chaix, Louise Deléger, Robert Bossy and Claire Nédellec

firstname.lastname@inra.fr

Microbial ecosystems

Which microbes live in an environment?



Properties of environment? Microbial Interaction?





Crossing between microbial species and habitats

Difficulty : Highly variable forms in text or genomics database (GOLD, SRA, GenBank)

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 <u>Enterococcus</u> and Lactobacillus and coliform bacteria were isolated from Dutch-type semi-hard cheese"



High incidence of Listeria monocytogenes in European red smear cheese

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

JGI 🎇 📕	G@LD		JGI H	IOME LOG IN		
(+••) (GENOMES ONLINE DA	TABASE				
ome Search Distrit	bution Graphs	Biogeographical Metadata	a Statistics GOLI	D Usage Policy Team Help	News	
Studies	28 817	Biosample Information	Biosample Source	Environmental Metadata		
Biosamples 🖯	23 362					
Sequencing Projects 🖯	139 492	BIOSAMPLE NAME				
nalysis Projects 🖯			Gb0125061			
Irganisms	275 649	Biosample Name 0		Microbial communities from thermized milk in Italy - ThermizedMilk.RNAseg.exp		
-		Other Names				
				Thermized milk		
		Community (microbial community		
		Community O		Ealy		
				ThermizedMik.RNAseg.exp1		
		Identifier 🖯				
		Add Date		2015-12-18		
		Last Modified By Tatiparthi Reddy on 2016-08-26				
		BIOSAMPLE DESCRIPTION				
		Biosample Description				
		Biosample Information				
		Biosample Information Vi	sibility 🖯	Yes		
		Ecosystem		Engineered		
		Ecosystem Category		Food production		
		Ecosystem Type		Dairy products		
		Ecosystem Subtype		Unclassified		
		Specific Ecosystem Unclassified BIOSAMPLE COMPOSITION				
		Number of Studies Number of Seg Projects		1		
		Number of Seq Projects Number of Analysis Project		1		
		Number of Related Biosan		26		

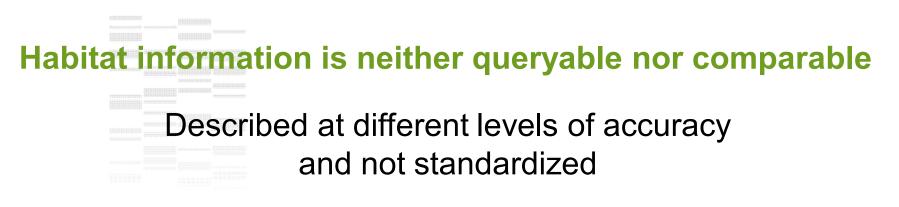
e.q.

- Artisanal cheeses from Tucuman
- ➤Dairy cheese

Caciocavallo cheese in Italy







What is the cheese microflora?

"Geotrichum candidum strains isolated from a <u>traditional Spanish goats'</u> <u>milk cheese</u>." *"Escherichia coli* O157:H7 isolated from raw beef, <u>soft cheese</u> and vegetables in Lima" *"Microbial ecology of <u>Gorgonzola</u> rinds and occurrence of different biotypes of <u>Listeria monocytogenes</u>."*



Classic search engine query

Google

cheese microbe

Scholar

Environ 25 800 résultats (0,06 s)

POST-PASTEURIAN CULTURES: The Microbiopolitics of Raw-Milk **Cheese** in the United States

H Paxson - Cultural Anthropology, 2008 - Wiley Online Library

... War II servicemen, the FDA now directs its sternest warning about **cheese**-residing **microbes** at pregnant ... 8 The fear is Listeria monocytogenes, the **microbe** behind listeriosis, which has been linked to ... that the category "soft" is neither self-evident nor used in the **cheese** world but ... Cité 166 fois Autres articles Les 15 versions Citer Enregistrer

The perils and promises of microbial abundance: Novel natures and model ecosystems, from artisanal **cheese** to alien seas H Paxson, S Helmreich - Social Studies of Science, 2014 - journals.sagepub.com Cité 43 fois Autres articles Les 7 versions Citer Enregistrer

Molecular approaches to analysing the microbial composition of raw milk and raw milk **cheese**

L Quigley, <u>O O'Sullivan</u>, TP Beresford, <u>RP Ross</u>... - International journal of ..., 2011 - Elsevier ... milk/cheese; Microbial composition; Culture-independent microbiology. 1. Introduction. Raw milk is known to harbour a complex microbial community. Indeed the high nutritional value of this food, its high water content and near neutral pH allows the growth of many **microbes** (... Cité 102 fois Autres articles Les 8 versions Citer Enregistrer

Food commensal **microbes** as a potentially important avenue in transmitting antibiotic resistance genes

HH Wang, M Manuzon, M Lehman... - FEMS ..., 2006 - femsle.oxfordjournals.org ... 1. Twenty out of the 23 **cheese** samples analyzed contained Tet r and/or Em r **microbes** ranging from 10 2 to 10 7 CFU g –1 of food, which are equivalent to 10 3 –10 8 CFU ART **microbes** per slice of **cheese** (about 20 g). In general, the number of Tet r **microbes** was greater in ... Cité 126 fois Autres articles Les 8 versions Citer Enregistrer

[LIVRE] Cheese and microbes

CW Donnelly - 2014 - books.google.com

A scientific overview of the association of **microbes** with **cheese**, through the lens of select **cheese** varieties that result due to surface mold ripening, internal mold ripening, rind washing, cave aging, or surface smear rind development. Over the past decade, there has Cité 11 fois Autres articles Les 8 versions Citer Enregistrer

The query matches "cheese" and "microbe" **but not** "Camembert", "Roquefort" or "Listeria monocytogenes"



We propose a semantic search engine dedicated to microbial biodiversity in food.







Semantic search engine of microbial habitat in food

Interpretation of the query

Aspergillus (taxon)

Aspergillus :

Synonyms (3)
 Aspergillus
 Aspergilli
 Petromyces

Sub-concepts (50) Aspergillus aculeatus Aspergillus amstelodami Aspergillus clavatus Aspergillus ficuum



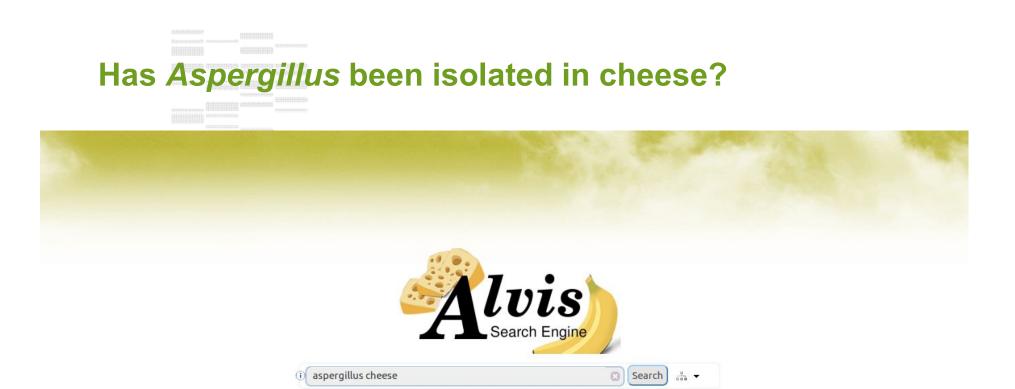
Cheese :

American cheese Cancoillotte Crème de Brie de Meaux Kiri The Laughing Cow brocciu caciocavallo chhena cottage cheese cream cheese

http://bibliome.jouy.inra.fr/demo/food/alvisir/webapi/search Mini-link: https://frama.link/AlvisFood











Results of the query: aspergillus cheese

			In I is a page 1 of 6 b is 0 In I is a page 1 of 6 b is 0 In I is a page 1 of 6 b is 0 In I is a page 1 of 6 b is 0	Aspergillus (taxon)
			10 1 k 4 Page 1 of 6 k K C [1 to 10 of 58]	Synonyms (3)
Microorga	nisms			Aspergillus
	10	20 All	Impact of microencapsulated peptidase (Aspergillus oryzae) on [cheddar] cheese proteolysis and its biologically active peptide profile. 20	Aspergilli
cet value	freq.	doc.	2011 Protein and peptide letters	Petromyces
spergillus	25	19		Sub-concepts (50)
ngi	34	19	Abstract We investigated the delivery of calcium-alginate encapsulated peptidase (Flavourzyme(®), Aspergillus oryzae) on proteolysis of (Cheddar cheese). Physical and	(There are the barrent) (
nicillium	26	17	chemical characteristics such as moisture, pH and fat content were measured, and no differences were found between control and [experimental cheese at day] 0. SDS-PAGE	cheese (habitat)
pergillus niger	30	12	analysis clearly showed that proteolysis of α and k casein was significantly accelerated after three months of maturity in the experimental cheese . A large number of low	Synonyms (1)
nicillium roqueforti	17	11		Sub-concepts (73)
pergillus versicolor	12	10	molecular weight peptides were found in the water soluble fraction of the experimental cheeses and some of these peptides were new. N-terminal amino acid sequence	fermented cheese
ccharomyces cerevisia	24	10	analysis identified these as P(1), Leu-Thu-Glu; P(3), Asp-Val-Pro-Ser-Glu) and relatively abundant stable peptides P(2), P(4), Arg-Pro-Lys-His-Pro-Ile; P(5), Arg-Pro-Lys-His-Pro-Ile;	fresh cheese
pergillus flavus	19	10	Lys and P(6). These peptides were mainly originated from αs1-CN and β-CN. Three of the identified peptides (P(1), P(2), P(3) and P(4)) are known to biologically active and	processed cheese
cteria	18	7	P(1) and P(3) were only present in experimental cheese suggesting that experimental cheese has improved health benefits.	ripened cheese
nicillium commune	6	6		streched curd chees
	25.5		»	whey cheese
Habita	ts			American cheese
	10	20 All		Cancoillotte
t value	freq.	doc.	The effect of Egyptian honeybee propolis on the growth of Aspergillus versicolor and sterigmatocystin biosynthesis in Ras cheese. 20	Crème de Brie de Me
ese	189	52	2007 The Journal of dairy research	Kiri
ng organism	33	19		The Laughing Cow
d	32	17	Abstract Propolis is a natural product collected by honeybee workers. The product was tested for its antifungal effect against Aspergillus versicolor ATCC 12996 as well as	brocciu
perimental medium	29	13	biosynthesis of sterigmatocystin during ripening of (Egyptian Ras cheese). The use of different concentrations of aqueous propolis extract 250, 500 and 1000 part per million	caciocavallo chhena
I	12	9	(ppm) on the cheese surface was investigated. Mould growth and toxin production were completely inhibited at the highest concentration 1000 ppm, while the lower	cottage cheese
gi and related product	13	9	concentrations exhibited definite fungistatic activity during 90 days of ripening. Control cheese demonstrated that the amount of sterigmatocystin produced was	cream cheese
k	15	7	proportional to the growth of Asp. versicolor during three months of ripening. It could be concluded that propolis concentration of 1000 ppm could prevent mould growth	faisselle
ry industry	13	6		feta
ist	7	6	and sterigmatocystin production in Ras cheese. The economic as well as the public health importance of propolis as a natural preservative in cheese manufacture is	fromage blanc
t and primary derivati	8	6	discussed.	hard cheese
				metata ayib
Journa	Is		»	mozzarella
	10	20 All		quark
t value	frea.	doc.	Effect of fermentation conditions on the production of citric acid from cheese whey by Aspergillus niger. 20	queso blanco
ernational Journal Of F	8	8		queso fresco ricotta
mal Of Dairy Sci	7	7	1996 International journal of food microbiology	schabziger
rnal Of Food Protect	3	3	Abstract The effect of pH volue, methanol, and self concentration on the production of sitis and from the volution of the state of the available set	semi soft cheese
lied Microbiology And	3	3	Abstract The effect of pH value, methanol, and salt concentration on the production of citric acid from cheese whey by two strains of Aspergillus niger, i.e. CAIM 111 and	soft cheese
tschrift Fur Lebensmit	2	2	CAIM 167, was investigated. Lactose concentration, utilized lactose, citric acid concentration, conversion coefficient of lactose to citric acid, and mycelial dry weight were	velveeta
Journal Of Dairy Rese	2	2	measured during the fermentation process. The maximum citric acid concentration (1.06 and 0.82 g/l), and conversion coefficient (5.58 and 7.45%) were obtained at pH 3.5	Beaufort
e Nahrung	2	2	after 9 days of fermentation for A. niger CAIM 111 and A. niger CAIM 167, respectively. The presence of 4% (v/v) methanol in the fermentation medium increased the	Brick





Result of the query: aspergillus cheese



Search 🚠

Impact of microencapsulated peptidase (Aspergillus oryzae) on cheddar cheese proteolysis and its biologically active peptide profile. 20

2011 Protein and peptide letters

lvis

Abstract We investigated the delivery of calcium-alginate encapsulated peptidase (Flavourzyme(\circledast), **Aspergillus oryzae**) on proteolysis of **Cheddar** cheese. Physical and chemical characteristics such as moisture, pH and fat content were measured, and no differences were found between control and **experimental cheese** at day 0. SDS-PAGE analysis clearly showed that proteolysis of α and k casein was significantly accelerated after three months of maturity in the **experimental cheese**. A large number of low molecular weight peptides were found in the water soluble fraction of the **experimental cheeses** and some of these peptides were new. N-terminal amino acid sequence analysis identified these as P(1), Leu-Thu-Glu; P(3), Asp-Val-Pro-Ser-Glu) and relatively abundant stable peptides P(2), P(4), Arg-Pro-Lys-His-Pro-Ile; P(5), Arg-Pro-Lys-His-Pro-Ile-Lys and P(6). These peptides were mainly originated from α s1-CN and β -CN. Three of the identified peptides (P(1), P(2), P(3) and P(4)) are known to biologically active and P(1) and P(3) were only present in **experimental cheese** suggesting that **experimental cheese** has improved health benefits.

>>

The effect of Egyptian honeybee propolis on the growth of Aspergillus versicolor and sterigmatocystin biosynthesis in Ras cheese. 20

2007 The Journal of dairy research

Abstract Propolis is a natural product collected by honeybee workers. The product was tested for its antifungal effect against **Aspergillus versicolor** ATCC 12996 as well as biosynthesis of sterigmatocystin during ripening of **Egyptian Ras cheese**. The use of different concentrations of aqueous propolis extract 250, 500 and 1000 part per million (ppm) on the **cheese surface** was investigated. Mould growth and toxin production were completely inhibited at the highest concentration 1000 ppm, while the lower concentrations exhibited definite fungistatic activity during 90 days of ripening. **Control cheese** demonstrated that the amount of sterigmatocystin produced was proportional to the growth of Asp. versicolor during three months of ripening. It could be concluded that propolis concentration of 1000 ppm could prevent mould growth and sterigmatocystin production in **Ras cheese**. The economic as well as the public health importance of propolis as a natural preservative in **cheese** manufacture is discussed.

>>











Result of the query: aspergillus ~livesin cheese

i aspergillus ~livesin cheese

Search

The effect of Egyptian honeybee propolis on the growth of Aspergillus versicolor and sterigmatocystin biosynthesis in Ras cheese.

2007 The Journal of dairy research

lvis

Abstract Propolis is a natural product collected by honeybee workers. The product was tested for its antifungal effect against Aspergillus versicolor ATCC 12996 as well as biosynthesis of sterigmatocystin during ripening of Egyptian Ras cheese. The use of different concentrations of aqueous propolis extract 250, 500 and 1000 part per million (ppm) on the cheese surface was investigated. Mould growth and toxin production were completely inhibited at the highest concentration 1000 ppm, while the lower concentrations exhibited definite fungistatic activity during 90 days of ripening. Control cheese demonstrated that the amount of sterigmatocystin produced was proportional to the growth of Asp. versicolor during three months of ripening. It could be concluded that propolis concentration of 1000 ppm could prevent mould growth and sterigmatocystin production in Ras cheese. The economic as well as the public health importance of propolis as a natural preservative in cheese manufacture is discussed.

>>

Effect of Zataria multiflora Boiss. essential oil on growth and aflatoxin formation by Aspergillus flavus in culture media and cheese. 1.414213

2009 Food and chemical toxicology : an international journal published for the British Industrial Biological Research Association

Abstract The effect of Zataria multiflora Boiss. essential oil (EO) against growth, spore production and aflatoxin formation by Aspergillus flavus (ATCC 15546) was investigated in synthetic media as well as Iranian ultra-filtered white cheese in brine. EO effectively inhibited radial growth and spore production on potato dextrose agar (PDA) in a dosedependent manner. At 200 ppm, the radial growth and sporulation reduced by 79.4% and 92.5%, respectively. The growth was completely prevented at EO400 ppm on PDA, and minimum fungicidal concentration (MFC) of the oil was estimated at 1000 ppm. The oil also significantly suppressed mycelial growth and aflatoxin synthesis in broth medium at all concentrations tested (P<0.05). At 150 ppm of EO, the mycelial growth and aflatoxin accumulation reduced by 90% and 99.4%, respectively. The EO at all concentrations tested, had an inhibitory effect against radial fungal growth and aflatoxin production by **A. flavus** in **Cheese**. However, no concentration of EO examined was able to completely inhibit the growth and aflatoxin production in **Cheese**. The results suggested the potential substitution of the antifungal chemicals by this EO as a natural inhibitor to control the growth of molds in foods such as **Cheese**.

>>





Behind the AlvisFood Search Engine

≻Our approach is to extract from text

- \rightarrow "Microbe" and "Habitat" concepts
- \rightarrow Links between them

≻We use

 \rightarrow AlvisNLP: Methods and tools for automatic extraction and analysis of biological text (*i.e.* Text Mining and Natural Language Processing)

→Machine learning methods trained with examples from microbiological and food domain experts

 \rightarrow Internal and external resources





Microbial entity detection

NCBI taxonomy

• Fungi (fungi) Click on organism name to get more information.

- Blastocladiomycota
 - Blastocladiomycetes
 - Blastocladiales
 - <u>Blastocladiales incertae sedis</u>
 - environmental samples
 - <u>uncultured Blastocladiomycota</u>
- <u>Chytridiomycota</u>
 - Chytridiomycetes
 - <u>Chytridiales</u>
 - <u>Cladochytriales</u>
 - Gromochytriales
 - Lobulomycetales
 - Mesochytriales
 - Polychytriales
 - <u>Rhizophlyctidales</u>
 - Rhizophydiales
 - <u>Spizellomycetales</u>
 - <u>unclassified Chytridiomycetes</u>
 - <u>Chytridiomycetes incertae sedis</u>
 - environmental samples
 - Monoblepharidomycetes
 - Monoblepharidales
 - <u>unclassified Monoblepharidomycetes</u>





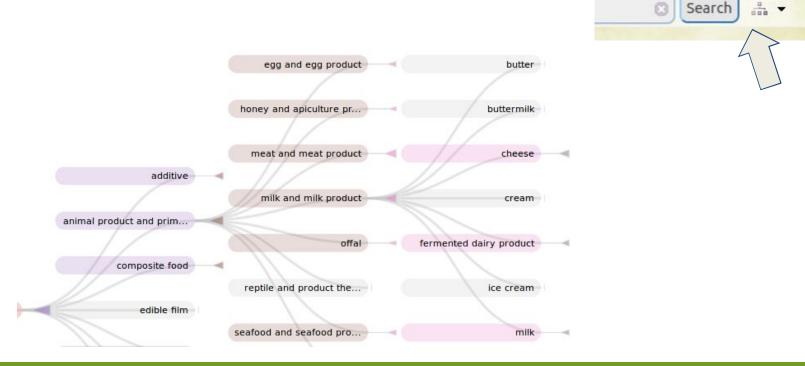
Habitat entity detection

Detection in text of nominal or adjectival groups

Categorization of these groups with the Ontobiotope ontology

 \rightarrow Formal and structured representation of microbial habitats

 \rightarrow Partially reused in AlvisFoodSE







Food sub-categories of Ontobiotope ontology

➢ From the EFSA classification

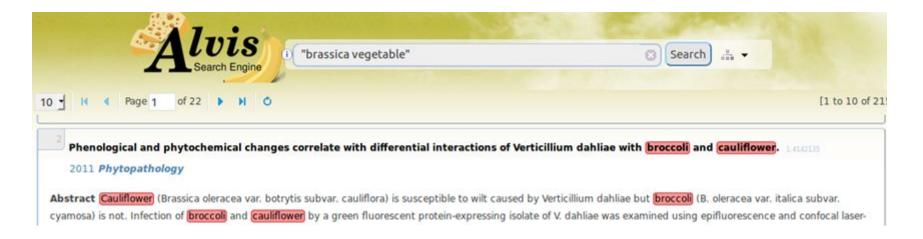
>Enrichment by microbial and food domains experts

≻Formal indication that "Roquefort" is a "Cheese"

 \rightarrow allows semantic search

➢Our automatic AlvisNLP tools link groups of words from the text to an Ontobiotope category

 \rightarrow achieve normalisation





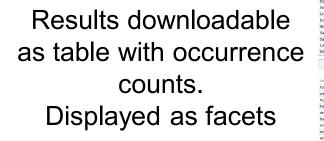
Relationship between Microbe and Habitat

>Extraction of ~livesin relationship

➢Hard problems in automatic language processing and artificial intelligence

Achieved by machine learning methods trained with annotated examples

> What are the taxa living in food? A query : {taxon}* ~livesin food









<u>http://bibliome.jouy.inra.fr/demo/food/alvisir/webapi/search</u> Mini-link: <u>https://frama.link/AlvisFood</u>

➢Our tools are pioneers in the field of text-mining for microbial biodiversity

 \succ Bibliome is a research team so:

 \rightarrow If you use AlvisFoodSE for your research, please cite us \rightarrow If you see an error, please send us an email, this will help us to improve our tools







<u>"Byssochlamys fulva</u> and <u>Neosartorya fischeri</u> are <u>heat-resistant fungi</u> which are a concern to food industries"

Automatic detection of microbial phenotypes

i.e. halophile, thermophile, phototroph ...

Acknowledgments

INRA Ontobiotope and Florilège working groups Food Microbiome project







Thank you for your attention





Text mining tools for extracting information about microbial biodiversity in food

Estelle Chaix, Louise Deléger, Robert Bossy and Claire Nédellec

Bibliome Team MaIAGE, INRA, Université Paris-Saclay, 78350 Jouy-en-Josas, France

firstname.lastname@inra.fr

http://bibliome.jouy.inra.fr/demo/food/alvisir/webapi/search Mini-link: https://frama.link/AlvisFood



