DOCTORAL THESIS PROJECT PROPOSAL 2021-2024

To submit no later than 10 December 2021

to Sylvie Fayard, E-mail: sylvie.fayard@univ-avignon.fr

Appel ciblé (merci de cocher la case correspondante) :

X Contrat doctoral 100% EUR Implanteus

□ Contrat doctoral 50% INRAE – 50% EUR Implanteus

Supervision

Thesis supervisor (HDR*, affiliated to DS536) ***Name & Research unit*** ***Number of doctoral theses currently supervised*****	Gérald CULIOLI (IMBE-IRPNC) (0,33 thesis currently supervised; started Fall 2019, Univ. Toulon)
Co-supervisor (HDR) ***Name & Research unit***	Olivier DANGLES (SQPOV-MicroNut) Carole MATHE DE SOUZA (IMBE-IRPNC)
Other researchers involved (Non-HDR) ***Names & Research units***	Olivier CHEVALLIER (PF 3A)

** Accreditation for Research Supervision*

*** According to the rules of Doctoral School 536 (DS536), researchers with a HDR cannot operate full supervision of more than 3 doctoral students simultaneously.*

Research units involved	**IMBE-IRPNC and SQPOV-MicroNut, AU (+ PF 3A)**
Title of the thesis project*	**New analytical developments for metabolomics and annotation of dyeing plant extracts: Applications in a cultural heritage context**

Summary *(10 – 15 lines)*

Since the dawn of time, natural dyes have been constantly used by humans until the advent of synthetic coloring molecules. In the field of cultural heritage and conservation practices, the detection of natural dyes in a wide range of cultural objects and archaeological samples could be of great help in understanding societal and scientific developments in various cultures and historical periods. In this context, the chemical characterization of natural dyes requires the development of new analytical methods and a better understanding of their degradation mechanisms.

This PhD project aims at developing new LC-MS-based untargeted metabolomics approaches using original molecular networking and *in silico* metabolization tools for annotation purposes. These analytical methods will be applied to a selection of natural extracts obtained from dye plants in order to identify taxonomic chemomarkers but also degradation compounds related to ancient dyeing techniques. Then, historical samples will be studied in order to determine the presence of such chemomarkers and thus to identify the dyeing plants and the techniques used in their production.

******The title will be published on the websites of the Doctoral School and Graduate School*

Description of the project (4 pages maximum)

Scientific and socio-economic stakes

The use of natural dyes by humans dates back to prehistoric times with the use of charcoal and mineral oxides. The first traces of plant coloring material have been assigned to indigo (2500 and 2000 BC. in India). In Europe, indigo has also been found in sites dated to the Bronze Age (1800-1000 BC.). The use and mastery of plant dyes go from Antiquity and reach a peak at the end of the nineteenth century. For example, traces of organic plant dyes can be found on Greek frescoes or on illuminations of medieval manuscripts. Until the end of the nineteenth century, plant dyes were one of the most important sources of coloring matter together with animal-derived dyes, such as purple (sea snail), kermes (insect), cochineal (insect), as well as other natural materials of mineral origin (Bechtold & Mussak 2009).

At the regional and local level, many dyeing plants have been widely cultivated in Provence, and more precisely in the Vaucluse: madder (*Rubia tinctorum*) in Sorgues and its surroundings or pastel (*Isatis tinctoria*) in the Cavaillon region (Peeters 1975). In addition, the city of Avignon was an important trading place of berries, known as "seeds of Avignon", from different buckthorns (*Rhamnus* spp.) from which a yellow matter was extracted.

In recent decades due to climate change and environmental problems related to polluting industries, consumers are increasingly sensitive to more sustainable production processes. Thus, the textile industry, which involves the use of toxic organic products (*e.g.* benzene, aniline, quinoline, dioxins) for dyes, is widely blamed (Khattab *et al.* 2020). The return to natural dyes is therefore in vogue with a production mainly from dyeing plants or co-products of the agri-food industry, but also through microbial biotechnology.

In this context, all the actors in this sector need powerful analytical tools in order:

(i) to verify the quality and to evaluate the traceability of the natural resources via a chemical inventory, as exhaustive as possible, of natural extracts from dyeing plants (*e.g.* indigo, madder, buckthorn, safflower, alkanet),

(ii) to validate the potential in these fields of natural extracts (which offers an incomparable range of natural colors) by trying to better understand the role of specific molecules in color expression,

(iii) to predict the potential of these extracts by studying the mechanisms of degradation of coloring molecules and by trying to better understand the strategies put in place by plants to stabilize them.

The aim of this thesis project is to apply to a panel of dyeing plants, selected on the basis of the chemical families they produced, original analytical protocols for the <u>most comprehensive chemical characterization of dyes extracts via untargeted metabolomics</u> approaches.

Metabolomics is the latest in the so-called "omics" sciences (*e.g.* genomics, transcriptomics, proteomics) that are intended for the systematic analysis of Life at the molecular level (Tautenhahn *et al.* 2012). Over the past few years, metabolomics has experienced a very significant boom with an exponential number of published scientific works and applications in fields as varied as medical and pharmaceutical sciences, agri-food, environmental sciences and ecology.

In the context of a metabolomic study, the main objective is to decipher the “metabolome” (the set of organic compounds of a non-polymeric nature quantifiable within an organism or more generally in a biological sample) and to evaluate and interpret its variations. The analysis

of the metabolome is therefore very complementary to those from other "omics" sciences as part of a global approach to biological systems (Patti *et al.* 2012). At present, untargeted metabolomic studies based on LC-MS (liquid chromatography coupled with mass spectrometry) are the most widely used, but the identification (the term "annotation" is often used for this purpose) of metabolites is generally the limiting factor in the interpretation of the resulting data (Chaleckis *et al.* 2019). Indeed, the diversity of metabolism is vast by nature and in a large part of metabolomic studies, the metabolites involved in discrimination between groups of samples often remain unidentified, thus limiting the biological interpretation of the results (Vaniya & Fiehn 2015).

In recent years, the LC-MS based molecular networking approach has emerged as a new tool to optimize the annotation process (Aksenov *et al.* 2021). Based on the idea that MS/MS fragmentation models are similar for molecules sharing similar chemical structures, the Global Natural Products Social Molecular Networking (GNPS) (http://gnps.ucsd.edu) web platform enables the creation of molecular networks in which compounds are related according to their molecular similarity (Aron *et al.* 2020; Wang *et al.* 2016). Currently, new and very original tools are being developed to increase the capabilities of molecular networks. For example, such tools allowed the extraction of recurring spectral patterns in complex datasets or metabolite "anticipation" using the *in silico* metabolization of identified target compounds (See Beniddir *et al.* 2021 for a review).

In conclusion, this PhD project will aim to develop the use of original analytical tools dedicated to the full chemical characterization of natural matrices through LC-MS-based metabolomics. These tools will be applied to the study of extracts obtained from a large panel of dyeing plants in order to: (i) evaluate the chemodiversity of such plants and to find robust taxonomical chemomarkers, (ii) evaluate the stability/degradation of coloring agents in these extracts, and finally (iii) help to recover ancient dyeing recipes and to decipher the natural dyes used in a selection of cultural heritage samples.

Research questions

In the current context of exponential development of metabolomics, there is an urgent need to integrate complementary analytical skills in order to improve our ability to decipher the in-depth chemical composition of non-model plants. Within the framework of research works dedicated to the study of natural dyes and their ancestral uses, the following questions can be raised:

- Can robust and innovative metabolomics approaches enable chemical discrimination of a wide selection of dyeing plants?
- Is it possible to use novel analytical tools (molecular networking, *in silico* metabolization) to annotate specific chemomarkers (*e.g.* taxonomic, degradation) in these natural matrices?
- Can these analytical methods and tools improve our knowledge of the stability and degradation mechanisms of coloring molecules from selected plant extracts, and this in relation with ancestral techniques used in the field of natural dyes.
- Can these analytical methods be applied to cultural heritage objects? Can such chemical characterization of natural dyes facilitate the restoration of heritage samples and improve our knowledge of ancestral recipes in this field? Can these recipes be used to develop new natural and effective alternatives to the synthetic dyes currently used?

Research program

In this context, this PhD thesis project will aim to:

1) establish a complete metabolomics workflow by LC-HRMS involving the implementation of a very original annotation step using molecular networks and to apply it to a large panel of dyeing plants representative of the chemical diversity of coloring molecules (*e.g.* flavonoids, anthocyanins, anthraquinones, indigoids…). This workflow will be developed in close collaboration with the metabolomics part of the 3A platform. The IMBE-IRPNC lab already holds many reference samples related to natural dyes (plant extracts, chemical standards). A collaboration will be also implemented with the "Jardin des plantes tinctoriales" in Lauris (Vaucluse, France), which has more than 250 different species of dyeing plants and the association "Couleurs Garance", which aims to promote plant dyes and their applications.

2) develop a methodology for anticipating the behavior of natural colored substances by *in silico* metabolism and to apply it to the study of the stability of selected dyes extracts and molecules. For that purpose, different degradation mechanisms (*e.g.* pH, T) and interactions (*e.g.* co-pigments, metal ions) will be considered in relation to the ancient applications of these dyes extracts.

3) search for characteristic dyes chemomarkers in the field of cultural heritage, as the identification of the natural dyes contained in historical samples could allow the deciphering of the ancestral dyeing recipes. Moreover, such studies are necessary for the documentation of artworks and may be decisive for the development of effective and appropriate conservation strategies. Various samples (textiles, votive objects...) will be selected from the current regional (*e.g.* Musée départemental d'Arles), national (*e.g.* Musée du Quai Branly, Paris) and international collaborations (*e.g.* Bénin, Amérique centrale) of the IMBE-IRPNC lab.

Provisional calendar

	Year 1		Year 2		Year 3	
Timeline (months)	**1-6**	**7-12**	**13-18**	**19-24**	**25-30**	**31-36**
Bibliographic survey						
Metabolomics approach development & Annotation tools implementation						
Analysis of a large panel of dyes extracts & Selection of a limited number of pertinent dyes extracts						
Study of stability & degradation of selected dyes extracts with respect to ancient recipes						
Characterization of natural dyes & ancient techniques in cultural heritage samples						
Data analysis Writing of articles & thesis Ms						

Suitability for the training program of the graduate school. Project multidisciplinarity

This PhD project falls under the Implanteus thematic axes I (*Crop production in a Mediterranean context*), II (*Preservation, extraction, processing and formulation of plant matter***)** and IV (*Interdisciplinary teaching* and more particularly: *Data processing, Statistics* and *Metabolomics*).

This project mainly focused on plant chemistry is clearly multidisciplinary, as it will require additional skills in various scientific areas such as plant biology, statistics, data handling… It will also benefit from a societal opening because of the ancestral uses of natural dyes in Provence, and particularly in Vaucluse.

Competences expected to be acquired by the doctoral student

The PhD student will acquire extensive practical and theoretical experience in:

- liquid chromatography and high-resolution mass spectrometry,
- structural characterization of natural products,
- metabolomics, data analysis and chemometrics.

The PhD student will acquire in-depth knowledge in the field of natural plant dyes (plant physiology, chemistry, dyeing processes) and in cultural heritage chemistry. He will learn to answer scientific questions using a variety of approaches and methods.
The PhD student will have to develop ability to work in team, rigor, enthusiasm, organization together with strong verbal and written communication skills.

We look for a talented student with a completed MSc. Degreein (bio)chemistry.The PhD student has a demonstrated track record of excellence in analytical (bio)chemistry and has extensive practical and theoretical experience inat least two of the following three areas:-liquid chromatography and high-resolution mass spectrometry,-structural characterization of natural products,-metabolomics, data analysis and chemometrics.The PhD student will acquire in-depth knowledge in the field of natural plant dyes (plant physiology, chemistry, dyeing processes) and in cultural heritage chemistry. He/shewill learn to answer scientific questions using a variety of approaches and methods.The PhD student will have to develop ability to work in team, rigor, enthusiasm, organization,together with strong oral and written communication skills.

Scientific partners, partners from the private sector / national and international

This thesis project will be led by researchers from 2 UMR teams involved in Implanteus (IMBE-IRPNC and SQPOV-MicroNut) who will each bring essential expertise in:

1) Chemistry of natural products and metabolomics/molecular networking (*e.g.* Cariot *et al.* 2021, Favre *et al.* 2017),
2) Structure/reactivity of polyphenols and food dyes (*e.g.* Trouillas *et al.* 2016, Fenger *et al.* 2021),
3) Cultural heritage chemistry and natural dyes used in this context (*e.g.* Cuoco *et al.* 2011, 2014).

The metabolomics part of this project will be carried out in close collaboration with the members of the 3A platform while the selection of dyeing plants and the dyeing processes will be discussed with members of the association “Couleur Garance”.
In the medium term, the analytical protocols developed and the knowledge accumulated during this PhD project could be of interest to local researchers (AU, INRAE center) working in the field of biomolecules, as well as to regional, national and international companies whose activities are related to the field of dyes and pigments.

Bibliographic references

Aksenov *et al.* (2021) "Auto-deconvolution and molecular networking of gas chromatography-mass spectrometry data" *Nat. Biotechnol.* 39: 169.
Aron *et al.* (2020) "Reproducible molecular networking of untargeted mass spectrometry data using GNPS" *Nat. Protoc.* 15: 1954.
Bechtold & Mussak (2009) *Handbook of Natural Colorants*. Chichester, John Wiley & Sons.
Beniddir *et al.* (2021) "Advances in decomposing complex metabolite mixtures using substructure- and network-based computational metabolomics approaches" *Nat. Prod. Rep.* 38: 1967.
Carriot *et al.* (2021) "Integration of LC/MS-based molecular networking and classical phytochemical approach allows in-depth annotation of the metabolome of non-model organisms - The case study of the brown seaweed *Taonia atomaria*" *Talanta* 225: 121925.
Chaleckis *et al.* (2019) "Challenges, progress and promises of metabolite annotation for LC–MS-based metabolomics" *Curr. Opin. Biotechnol.* 55: 44.
Cuoco *et al.* (2011) "Characterization of madder and garancine in historic French red materials by liquid chromatography-photodiode array detection" *J. Cult. Herit.* 12(1): 98.

Cuoco *et al.* (2014) "Liquid chromatographic analysis of flavonol compounds in green fruits of three *Rhamnus* species used in Stil de grain" *Microchem. J.* 115: 130-137.
Favre *et al.* (2017) "Discrimination of four marine biofilm-forming bacteria by LC–MS metabolomics and influence of culture parameters" *J. Proteom. Res.* 16(5): 1962.
Fenger *et al.* (2020) "The fate of acylated anthocyanins in mildly heated neutral solution" Dyes Pigm. 178: 108326.
Khattab *et al.* (2020) "Textile dyeing industry: environmental impacts and remediation" *Environ. Sci. Pollut. Res.* 27(4): 3803.
Patti *et al.* (2012) "Meta-analysis of untargeted metabolomic data from multiple profiling experiments" *Nat. Protoc.* 7: 508.
Peeters (1975) "Les plantes tinctoriales dans l'économie du Vaucluse au XIXe siècle" *Études rurales* 60: 41.
Tautenhahn *et al.* (2012) "An accelerated workflow for untargeted metabolomics using the METLIN database" *Nat. Biotechnol.* 30: 826.
Trouillas *et al.* (2016) "Stabilizing and modulating color by copigmentation: Insights from theory and experiment" *Chem. Rev.* 116(9): 4937.
Vaniya & Fiehn (2015) "Using fragmentation trees and mass spectral trees for identifying unknown compounds in metabolomics" *Trends Anal. Chem.* 69: 52.
Wang *et al.* (2016) "Sharing and community curation of mass spectrometry data with Global Natural Products Social molecular networking" *Nat. Biotechnol.* 34: 828.

Other information (1/2 page maximum)

e.g.: possibility of specific funding for research needs, possibility of mobility to other laboratories...

The two teams will guarantee financial support for the necessary experiments (laboratory's own fundings, ongoing projects) and will set up appropriate projects to promote the student's international mobility.
As one of the partners (IMBE-IRPNC) is a member of an ongoing EU research project (ending in April 2022, http://www.erasmusmundus-archmat.uevora.pt/overview.html) that aims to promote the training of young researchers and to establish a perennial European network in the field of archaeometry, the mobility of the PhD student will be facilitated and directed toward the appropriate lab/partner.

Opinion of the research unit director(s)

In the case where one of the units involved in the project has submitted several requests for EUR funding, please rank the requests.

I, the undersigned, Thierry Dutoit, research director at CNRS and codirector of UMR IMBE give a very favorable opinion to the thesis project presented by our colleagues Gérald Culioli, Olivier Dangles and Carole Mathe De Souza. This project entitled "*New analytical developments for metabolomics and annotation of dyeing plant extracts: Applications in a cultural heritage context*" aims at the chemical characterization of natural dyes which requires the development of new analytical methods and a better understanding of their degradation mechanisms. It is in the heart of the activities of the chemists of IRPNC team at IMBE in Avignon and may subsequently constitute in the future an interesting bridge between chemists and the other biologist and environmental members of this team. The group of co-supervisors of this thesis has all the skills and background to carry out this project that is also based on the platform (PF 3A- Olivier Chevallier).

For all these raisons, the IMBE direction strongly supports this project, which will strengthen the integration into our unit of Gérald Culioli, a recently recruited professor.

Thierry DUTOIT

Directeur de recherches CNRS

Directeur adjoint IMBE

The project is in full accordance with the expertise developed at SQPOV about the reactivity of phytochemicals in food environments that affects their nutritional, taste and coloring properties. A leading scientist of SQPOV participates in the supervision the project. The project also rests on the analytical facilities offered by platform 3A and is shared with another research unit with great interest in phytochemicals: this a great opportunity for establishing new collaborations.

Very positive opinion

Priority SQPOV 1/3