

---

## Plan Overview

*A Data Management Plan created using DMPonline*

**Title:** BIONIC - BIOSynthetic production of tobacco-free NICotine and related pyridine-based alkaloids

**Creator:** James Galman

**Principal Investigator:** James Galman

**Data Manager:** James Galman

**Affiliation:** University of Manchester

**Template:** BBSRC Template Customised By: University of Manchester

**ORCID ID:** 0000-0003-2076-2806

### Project abstract:

Cultivated tobacco (*Nicotiana tabacum*) is the most commonly grown and studied species of all plants in the genus *Nicotiana* noted for the production of the potent alkaloid nicotine. The widespread use of tobacco is due to the addictive effects of nicotine which can be extracted from tobacco plants but requires further downstream processing to obtain pharmaceutical grade quality (>95%) for smoke-free tobacco cessation products. Biocatalysis provides a sustainable and 'green' alternative to cost prohibitive chemical processes minimising the waste and chemical steps needed to achieve optically pure nicotine. To date, a clear understanding of the enzyme(s) involved in the later stages of nicotine biosynthesis remains elusive. A main goal of the project is focused on enzyme discovery from a sequence-based metagenomics approach to strategically mimic the biosynthesis of high-valued tobacco-free nicotine and other pyridine-related products. Moreover, the newly discovered enzyme(s) will create de novo pathways to functionalised pyridine building blocks such as agrochemical and pharmaceutical intermediates inaccessible in the current biocatalyst toolbox. Intensification of these enzymatic processes will also be investigated to demonstrate high performance and increased space-time yields to be implemented on an industrial scale.

**ID:** 58691

**Last modified:** 10-06-2020

### Copyright information:

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customise it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

# BIONIC - BIOSynthetic production of tobacco-free NICotine and related pyridine-based alkaloids

---

## Manchester Data Management Outline

### 1. Will this project be reviewed by any of the following bodies (please select all that apply)?

- Funder

### 2. Is The University of Manchester collaborating with other institutions on this project?

- Yes - Part of a collaboration and owning or handling data

The applicant will be collaborating with an Industrial Partner (Prozomix Ltd) an SME that will provide metagenomic libraries of enzymes and sharing protein sequence data.

### 3. What data will you use in this project (please select all that apply)?

- Acquire new data

Raw data will be collected, analysed and stored from wet lab experiments.

### 4. Where will the data be stored and backed-up during the project lifetime?

- University of Manchester Research Data Storage Service (Isilon)
- P Drive (postgraduate researchers and students only)

The University of Manchester's Research Data Management Service (RDMS), which satisfies the Research Council UK RDM guidelines, is eligible for research staff will be used in this project as it provides safe, resilient, secure and managed replicated storage of scientific data. This service allows researchers to manage, store, and curate the data, as well as to preserve it after the end of the project

### 5. If you will be using Research Data Storage, how much storage will you require?

- < 1 TB

### 6. Are you going to be working with a 3rd party data provider?

- No

### 7. How long do you intend to keep your data for after the end of your project (in years)?

- 5 - 10 years

According to University of Manchester's RDMS Policy, relevant data that are likely to have long term value, including data that demonstrate research findings or represent records of the University, will be preserved (5-10 years) and curated after project completion for as long as they remain of value.

**Questions about personal information**

Personal information, also known as personal data, relates to identifiable living individuals. Special category personal data is more sensitive information such as medical records, ethnic background, religious beliefs, political opinions, sexual orientation and criminal convictions or offences information. If you are not using personal data then you can skip the rest of this section.

Please note that in line with [data protection law](#) (the General Data Protection Regulation and Data Protection Act 2018), personal information should only be stored in an identifiable form for as long as is necessary for the project; it should be pseudonymised (partially de-identified) and/or anonymised (completely de-identified) as soon as practically possible. You must obtain the appropriate [ethical approval](#) in order to use identifiable personal data.

**8. What type of person identifying information will you be processing (please select all that apply)?**

- No sensitive or personal data

**9. Please briefly outline how you plan to store, protect and ensure confidentiality of the participants' information.**

Not applicable

**10. If you are storing personal information (including contact details) will you need to keep it beyond the end of the project?**

- Not applicable

**11. Will the participants' information (personal and/or sensitive) be shared with or accessed by anyone outside of the University of Manchester?**

- Not applicable

**12. If you will be sharing personal information outside of the University of Manchester will the individual or organisation you are sharing with be outside the EEA?**

- Not applicable

**13. Are you planning to use the personal information for future purposes such as research?**

- No

**14. Who will act as the data custodian for this study, and so be responsible for the information involved?**

Line Manager, Dr. Ian Rowles

**15. Please provide the date on which this plan was last reviewed (dd/mm/yyyy).**

2020-06-05

## Data areas and data types

### Outline the volume, type and content of data that will be generated e.g. experimental measurements, models, records and images

The analysis will generate raw mass spectrometry files, UV-HPLC traces. Data will be acquired and stored on University Managed PCs and Laptops with large storage space, older data being archived to removable large capacity storage. Archived media will be stored at a secure secondary location.

## Standards and metadata

### Outline the standards and methodologies that will be adopted for data collection and management, and why these have been selected

The Fellow and the collaborator (Prozomix) will be responsible for quality of the data acquisition and the data management, ensuring the collection methods employed, is reproducible, consistent and reliable through standardized operating protocols. Regular check-ups on both the methodology and data collection methods will be reviewed and assessed by experienced colleagues to ensure the highest standards are met. Detailed descriptions of the experimental procedure will be stored and disseminated through high-impact publications and conference presentations.

## Relationship to other data

### State the relationship to other data available in public repositories

Any novel DNA or protein sequence/structural data collected as part of this project will be deposited in public databases (e.g. Genbank, Protein Data Bank, etc.) to make the data available to the wider scientific community.

## Secondary Use

### Outline the further intended and/or foreseeable research uses for the completed dataset(s)

Experimental methods and protocols will be included in the main text of the publications or as supporting information. The Fellow is fully aware of the University's policy to subsidise or fully fund Open access articles through the depositing of 'just accepted articles' and the acknowledgement of research grants provided by the UK research councils (e.g. BBSRC). This is funded through the Gold Open Access Scheme ensuring his research is widely read as possible.

## Methods for data sharing

### Outline the planned mechanisms for making these data available, e.g. through deposition in existing public databases or on request, including access mechanisms where appropriate

All plasmids and strains constructed will be archived and made available to share with academics upon request after completing a Materials Transfer Agreement (MTA).

## Proprietary data

### **Outline any restrictions on data sharing due to the need to protect proprietary or patentable data**

In order to protect intellectual property and respect confidentiality with the industrial partner (Prozomix) some restrictions may be necessary, but they will be minimised as far as practicable.

## **Timeframes**

### **State the timescales for public release of data**

The data generated during the Fellowship will be released in accordance with established practices within the relevant research disciplines. Data that will be made available through scientific publication will be written up and submitted promptly.

## **Formats**

### **State the format of the final dataset**

Data from HPLC and GC data will be in .D and converted to .csv format for analysis, furthermore .xlsx format for Plate readers.