

#31

**COMPLETE**

**Collector:** Web Link 1 (Web Link)  
**Started:** Sunday, December 03, 2017 4:57:31 PM  
**Last Modified:** Monday, December 11, 2017 9:47:41 AM  
**Time Spent:** Over a day  
**IP Address:** 45.49.140.111

---

Page 1: PhD Project Proforma 2018

**Q1** Please provide project supervisors contact details. These may be used by potential applicants wishing to discuss the project with you.

Supervisor 1 name	<b>Ammar Al Chalabi</b>
Supervisor 1 telephone	<b>+44 20 7848 5192</b>
Supervisor 1 email	<b>ammar.al-chalabi@kcl.ac.uk</b>
Supervisor 1 organisation	<b>King's College London</b>
Supervisor 2 name	<b>Alfredo Iacoangeli</b>
Supervisor 2 telephone	<b>+44 743 488 56 40</b>
Supervisor 2 email	<b>alfredo.iacoangeli@kcl.ac.uk</b>
Supervisor 2 organisation	<b>King's College London</b>

---

**Q2** Any additional supervisors, contact details and organisation **Respondent skipped this question**

---

**Q3** Title of Research Project

A high throughput bioinformatics system for the analysis of ALS/MND multi-layer omics, environmental and clinical data

---

**Q4** Date of Supervisor Training?

05/02/16

---

**Q5** Name of BRC Cluster the project sits under **Cluster 3 - Data Analytics**

---

**Q6** Which BRC Themes does your project align with? **Genomic Medicine**  
Tick all that apply

---

### Q7 Abstract of Research Project (max 500 words)

Amyotrophic lateral sclerosis (ALS, motor neuron disease) is a devastating disease causing progressive paralysis until death from respiratory failure over two-three years. ALS kills 1 in every 300 people and is the commonest reason to seek assisted suicide. The factors increasing and decreasing risk, modulating phenotype, altering age of onset, or triggering disease remain largely unknown except for some genetic variations between individuals and a few environmental exposures, thus the need for further research is great. Several international initiatives are worldwide collecting multilayered omics, environmental and clinical data of tens of thousands of patients, in order to elucidate its genetic bases and reveal its interface with the environment. Two among the leading ones are Project Mine (<http://www.projectmine.com>) and STRENGTH ([http://www.neurodegenerationresearch.eu/fileadmin/Project\\_Fact\\_Sheets/PDFs/Risk\\_Factors/STRENGTH.pdf](http://www.neurodegenerationresearch.eu/fileadmin/Project_Fact_Sheets/PDFs/Risk_Factors/STRENGTH.pdf)). Project MinE (co-lead Al-Chalabi, data infrastructure lead Iacoangeli), whose aim it to elucidate the genetic basis of this disease, is the biggest one-disease whole-genome sequencing project in the world with more than 22000 whole genomes sequenced on the Illumina platform, and accompanying clinical data. The STRENGTH consortium is focusing on the interaction between the genetics and the environment, accompanying the genetic data with epigenetic and environmental data. The results of this large scale projects comprise hundreds of Tb of data, spread over the USA, Australia and countries in Europe and elsewhere, making analysis, sharing and handling very challenging. While tools to facilitate sharing of data as a single set are being developed (<https://github.com/KHP-Informatics/MND-DataManagementAnalysis-System>), analysis tools are needed to make the extraction, visualization and interpretation of such a variety of data effective and accessible to clinicians and scientists. This PhD is a collaboration between clinical neuroscientists and health informatics experts working on ALS. The candidate will explore state-of-the-art bioinformatics tools and analysis methods to identify the most effective interfaces and analysis techniques, and analyse the structure of the multi-omics, clinical and environmental data (not publicly available) made available by Project MinE and STRENGTH, to identify and overcome the challenges it poses. Through an iterative process with users, and by analysing the research questions most asked of sequencing data by clinicians and scientists, the candidate will develop new tools, ICT set-ups and interfaces, building on what works best. The project will use open source technologies and cloud computing facilities. This will be part of a powerful and useful system for genomics research with the aim of making its usage and findings accessible and useful for the medical clinic, harnessing the great potentialities of next generation sequencing and other omics data to the everyday scientist and clinicians.

---

### Q8 Please provide up to three key references relevant to proposed work

- 1) Al-Chalabi A., Van den Berg, L. H. and Veldink, J. (2017). Gene discovery in amyotrophic lateral sclerosis: implications for clinical management. *Nat. Rev. Neurol.* 13, 96-104. doi:10.1038/nrneurol.2016.182
  - 2) Project MinE: study design and pilot analyses of a large-scale whole-genome sequencing study in amyotrophic lateral sclerosis - Project MinE Consortium, et al bioRxiv 152553; doi: <https://doi.org/10.1101/152553>
- 

### Q9 Outline the Departmental/School research training and in-house research support in relation to this project

The BRC Health Informatics unit has a strong bioinformatics and software development component. The bioinformatics sub-unit runs bioinformatics services for researchers and courses both for several Master programmes and as short summer courses. The Software development sub-unit includes both senior and junior developers who can provide technical guidance and support for the other members of the team. HPC facility support and training will be provided by the SLAM-BRC Rosalind team which is also part of the Health Informatics unit. Training in genetics and neurology, as well as guidance during the whole project will be provided by the Clinical Neuroscience unit. This unit is lead by Ammar Al-Chalabi who is also head of the Clinical Neuroscience master programme.

---

## PhD Projects 2018

**Q10** Summarise the computational and experimental medicine component of the research, as well as the relevance of the project to the BRC Research Cluster(s)

The major component of this project is computational. The candidate will be developing their bioinformatics and software development skills working within the Health Informatics unit (Richard Dobson lead) of the KCL Institute of Psychiatry, Psychology and Neurology. The candidate will also work alongside the clinicians and geneticists working in the Clinical Neuroscience unit (Ammar Al-Chalabi lead) at the Maurice Wohl Clinical Neuroscience institute. The research genetics and medical output will be tested and validated in wet lab and clinic. The candidate would be able to follow and take part to these phase but this is not be mandatory. The project is therefore of clear relevance for the Precision Medicine and Data Analytics BRC Research Clusters

---

**Q11** Supervisor 1 details: You are first or co-supervisor for in each year 1-4

3

---

**Q12** Please provide a KCL budget code relevant to the project **Respondent skipped this question**

---

**Q13** Supervisor 1: Students you have been supervisor for who have completed since October 2002 (Number completed and number who have completed within 4 years)

0

---

**Q14** Supervisor 2 details: Number of current PhD Student you are first supervisor for in each year 1-4

0

---

**Q15** Supervisor 2: Students you have been supervisor for who have completed since October 2002 (Number completed and number who have completed within 4 years) **Respondent skipped this question**

---

**Q16** Please provide project supervisors contact details - which can be used by potential applicants wishing to discuss the projects

alfredo iacoangeli, email [alfredo.iacoangeli@kcl.ac.uk](mailto:alfredo.iacoangeli@kcl.ac.uk)  
ammar al chalabi , email [ammar.al-chalabi@kcl.ac.uk](mailto:ammar.al-chalabi@kcl.ac.uk)

---