# An Education Incubator to Engage Biomedical Researchers with HPC

Andrea Townsend-Nicholson

Faculty of Life Sciences Division of Biosciences Structural & Molecular Biology

### **Computational Biomedicine**

#### The Vision



#### The Reality

KEY: starting early to engage biomedical scientists - create fluency, expectation and support innovation and new discoveries in high performance computing

#### Many barriers to overcome

- Access to the infrastructure
- Training in its use
- Perceived difficulty
- Perceived demographic

### CompBioMed Education Aims

To bridge the computational-experimental divide by supporting and developing existing users

To create a new category of users of HPC ("future users") who will be fluent both computationally and experimentally

To provide 'tried and tested' educational templates and expertise for adoptions by other universities and institutions

# The Landscape in 2016

| Field        | HE First Degree<br>(%) | HE Postgraduate Degree | Archer Use |
|--------------|------------------------|------------------------|------------|
| Medical      | 33                     | 24                     | 0.2        |
| Biological   | 26                     | 22                     | 9.8        |
| Physical     | 15                     | 25                     | 60.7       |
| Mathematical | 5                      | 4                      | 9.5        |
| Engineering  | 15                     | 20                     | 18.6       |
| Other        | 6                      | 4                      | 1.2        |
| Total        | 100                    | 100                    | 100        |

© Andrea Townsend-Nicholson 2020

# The Teaching in CompBioMed1

- Life Sciences undergraduates (BSc, iBSc, MSci)
  - BIOC0023 Specialist Research Project in Metagenomics
  - BIOC0031 Metagenomics and Molecular Medicine
  - BIOC0002 The Basics for Molecular Biosciences (Key Skills)
  - BIOC0012 Research Project Foundations (Key Skills)
- Medical students
  - SSC334 From Skin to Metagenomics, Year 1, Year 2
  - Extracurricular projects in Metagenomics (Year 2)

\*All module resources are placed on GitHub, on the UCL institutional virtual learning environment and are published on the CompBioMed training portal

# Successful Engagement!

| Academic<br>year | Medical<br>students | Molecular<br>Biosciences<br>students | Total<br>number of<br>students | Core hours<br>consumed | Core hours<br>allocated | Fold difference<br>(consumed/allocated) |
|------------------|---------------------|--------------------------------------|--------------------------------|------------------------|-------------------------|---|
| 2015-2016        | 0                   | 20                                   | 20                             | 0 (local)              | 0                       | -                                       |
| 2016-2017        | 0                   | 29                                   | 29                             | 0 (local)              | 0                       | -                                       |
| 2017-2018        | 40                  | 85                                   | 125                            | 17,452                 | 10,500                  | 1.66                                    |
| 2018-2019        | 20                  | 99                                   | 119                            | 49,394                 | 10,900                  | 4.53                                    |
| 2019-2020        | 20                  | 83                                   | 103                            | 97,919*                | 10,830                  | 9.04*                                   |

# The Outcomes

#### **Teaching Outcomes**

- 100% success rate for students using HPC as part of their degree (2017-2020)
- Improved diversity: >50% female\* >40% BME
- Employability from embedding computation in the Molecular Biosciences Curriculum

#### **Research Outcomes**

- Expansion of the programme: EU funding to deliver medical student HPC teaching at UVA, UPF/BSC, UOXF, USFD (CompBioMed2)
- Engagement with UKRI and EC To provide resource to support training HPC users, including students (£13m)

### CompBioMed Education Aims

To bridge the computational-experimental divide by supporting and developing existing users

To create a new category of users of HPC ("future users") who will be fluent both computationally and experimentally

To provide 'tried and tested' educational templates and expertise for adoptions by other universities and institutions

#### Next Steps – CompBioMed2

- CompBioMed2 Education Training: Expanding from UCL to Oxford, Sheffield, UPF and UvA
- Adding new elements of computational biology to the university curriculum (improved digital skills, machine learning, MD simulations)
- Training early career and established researchers
- Enhancing the User Experience

# Expanding from UCL - SSC@USFD (2020-2021)

#### Phase 2A Research SSC Project Title:

From Skin to Metagenomics: Exploring Your Microbiome

#### Supervisors:

Dr Andrew Narracott (Infection, Immunity & Cardiovascular Diseases, University of Sheffield), Prof Andrea Townsend-Nicholson (Division of Biosciences, University College London), Dr Guillaume Hautbergue (Sheffield Institute for Translational Neuroscience, University of Sheffield)

#### Associate Partner: Alces Flight

### Training Early Career and Established Researchers



- Introduction to Microbiomes and 16S rDNA
- Introduction to NGS (Next Generation Sequencing)
- Workshop Getting to know your dataset
- Workshop Computing and High Performance Computing (HPC)
- Analysis1 (command line)
- Analysis2 (Jupyter notebooks)
- Analysis3 (in the cloud)
- Comparison of data and summary of findings

26 May 2020

### CompBioMed2 Education and Alces Flight

- Proof of Concept Research Collaboration Partnership
- Embed web-based HPC in the USFD SSC for September 2020
- Capability to grow beyond the project into other avenues of training in cloud HPC should this be within the interest of UCL and *CompBioMed2* (e.g. Biochemical Society Training Event)

#### CompBioMed2 Education and Dassault Systèmes

- A new SSC to be trialled at UCL will be co-delivered by Dassault Systèmes in 2020-2021
- It will provide n overview of *in-silico* methods for simulation using Finite Element (FE) analysis
- This SSC provides a very rare opportunity to acquire hands-on expertise in this fast growing cardiovascular area of Virtual Human Modelling using the class-leading FE Software Abaqus.

### Conclusions

- The CompBioMed education programme provides a unique opportunity for an innovation incubator that will provide new users with access to HPC applications and the capacity to co-create new HPC applications with this nascent community
- Engagement with industrial partners enables access to state of the art technologies and applications

### Acknowledgments

#### **CompBioMed Partners**

University College London - Michael Baron and Amanda Cain, Jim Naismith (external examiner), Greg Campbell and Paul Dilworth, David Gregory, Peter Coveney University of Sheffield – Guillaume Hautbergue and Andrew Narracott

University of Edinburgh - Gavin Pringle and Terry Sloan SURFsara - Marco Verdicchio Barcelona Supercomputing Centre - Mariano Vazquez, Ruth Aris and Alfonso Santiago

Dassault Systèmes – Clint Taylor-Davies Alces Flight – Cristin Merritt and Wil Mayers



26 May 2020

© Andrea Townsend-Nicholson 2020