

EMBL
Australia



Seminar with Dr Alex Bateman

UniProt: A universal hub of protein knowledge



Monday, 1 April 2019

10:00 - 11:00am



Physiology Seminar Room,
Room F122D
1F 26 Innovation Walk (13F),
Monash University, Clayton.

Join us for a presentation where Dr Bateman will discuss the UniProt KnowledgeBase, which has been providing data for protein sequence and function for over 30 years.

Abstract

I will discuss how we are dealing with the deluge of sequence data to organise it in ways to make our users lives easier and more scalable. I will show with examples how our expert curators gather knowledge from the literature, evaluate it, and add value through structuring it with a variety of ontologies.

Does expert curation scale with the vast biomedical literature? I will try to address this important point. Firstly, through recent experiments to estimate the fraction of the literature we curate and secondly through automated systems to annotate proteins that have never been experimentally characterised. I will discuss how UniProt helps to connect and integrate the worldwide community of biological resources and their data. I will conclude with some recent machine learning research work to investigate whether we can estimate the fraction of spurious or erroneous protein sequences that are in the database.

About Dr Bateman

Dr Alex Bateman leads the Protein sequence resources cluster that includes world leading databases including UniProt, Pfam and Rfam. As Head of Protein Sequence Resources at EMBL-EBI, Alex Bateman also has an important strategic leadership role in the InterPro and UniProtKB databases. He joined EMBL-EBI as the Head of Protein Sequence Resources in 2012. Prior to that, he managed the production of numerous popular biological databases such as Pfam and Rfam at the Sanger Institute. He was formerly a member of, and Chairman of the ISB's Executive Committee. He was also Executive Editor for Bioinformatics and Editor of NAR's database issue.