Institution: The University of Edinburgh

Unit of Assessment: UoA5: Biological Sciences

Title of case study:

01. Phylogenetic analysis software BEAST informs public health responses to infection

1. Summary of the impact

Impact: BEAST software has widespread applications with impacts on public health policy, service provision and awareness, and in other contexts such as commercial disputes and criminal cases.

Beneficiaries: Public agencies such as health bodies and criminal courts; ultimately, global and local populations subject to infectious disease epidemic and pandemic outbreaks in which BEAST is used to inform the response.

Significance and Reach: BEAST is critical software that has been used to understand the spread of and to inform the response to global pandemics such as H1N1 swine-flu. It is also used to determine disease origin and transmission issues in specific situations (e.g. in criminal cases). The reach of this software is therefore both global and local.

Attribution: Rambaut (UoE) co-led the phylogenetic research and developed BEAST with Drummond (Auckland, NZ). The subsequent epidemic and pandemic analyses were variously led by Rambaut and Pybus (Oxford) and by Ferguson (Imperial College London).

2. Underpinning research

Andrew Rambaut of UoE, with Alexei Drummond (Auckland, NZ), developed the widely-used software package, **BEAST** (Bayesian Evolutionary Analysis Sampling Trees) first published in 2007 [1] http://beast.bio.ed.ac.uk/.

BEAST provides a general framework for parameter estimation and hypothesis testing of evolutionary models from molecular sequence data. The research that led to publication of BEAST is based on the development of computational methods for the analysis of gene sequence data from infectious disease agents such as viruses and bacteria to investigate their evolutionary and epidemiological dynamics. Information about the timing and spatial location of the origins of human and animal epidemics can be traced and the rate of epidemiological spread can be measured from a relatively small sample of isolates from the population of infected individuals. The key insight is that many small intracellular pathogens evolve at a rate that is sufficiently rapid that the genetic mutations accumulated over the course of an epidemic record information about the epidemiological processes that underpin the disease.

BEAST is a Bayesian numerical estimation framework for evolutionary analysis with a particular emphasis on infectious disease (although it is widely used in other fields). It is a fast, flexible software architecture for Bayesian analysis of molecular sequences related by an evolutionary tree. A large number of popular stochastic models of sequence evolution are provided and tree-based models suitable for both within- and between-species sequence data are implemented. BEAST consists of object-oriented Java source code and uses Metropolis-Hastings Markov chain Monte Carlo Bayesian analysis. It provides models for DNA and protein sequence evolution, highly parametric coalescent analysis, relaxed clock phylogenetics, non-contemporaneous sequence data, statistical alignment and a wide range of options for prior distributions. BEAST source code is highly modular allowing other developers to easily add models or components. The components can be combined in almost unlimited combinations allowing the construction of complex models to describe particular infectious diseases or epidemics and to address particular questions. For example, it is possible to combine an epidemiologically informed model of the growth of new infections with a molecular evolutionary model of a viral genome and a spatial model of the movement of individuals amongst locations to reconstruct the time and location of the source of an epidemic.

The software has been developed with a powerful but easy to use interface to make it accessible to those working in public health. It is also provides a set of post-analysis visualization tools to allow both statistical and visual representation of the results. It is a free open-source program which now has a wider group of registered developers and contributors, overseen by Rambaut. The original BMC paper [1] is flagged as 'highly accessed' on BioMed Central.

UoE research outputs using BEAST include (i) understanding the global movement of influenza virus and the seeding of Northern and Southern hemisphere winter epidemics by strains in non-temperate areas [2], (ii) demonstrating the nosocomial nature of an outbreak of HIV and hepatitis C virus amongst children in a Libyan hospital (which exonerated the foreign medical workers who had been accused of deliberate infection and sentenced to death) [3]; and (iii) revealing the zoonotic origins and genetic architecture of the H1N1 influenza A pandemic virus [4]. This last example, in particular, demonstrates the power of the software because, by the time the virus had been discovered in California, it had had already spread extremely widely unnoticed and would have been impossible to trace by conventional epidemiological means. Genetic data combined with the statistical models implemented in BEAST based on data from the initial outbreak in Mexico allowed the reconstruction of the hidden early history of the pandemic. BEAST was used in the Science paper [5], which estimated the start of the outbreak by looking at the diversity in the genetic sequences of viral samples collected from confirmed cases, assuming that diversity accumulates according to a molecular clock model. Twenty-three complete publicly available hemagglutinin gene sequences from cases not linked in epidemiological clusters were analysed using BEAST. This gave estimates of the date of entry of the strain into the human population and its initial rate of spread in terms of the basic reproductive number, a key epidemiological property required to understand the likely effect of control measures and interventions [5].

BEAST is co-authored by Professor Andrew Rambaut (UoE, 2006-present) and A. Drummond (Bioinformatics Institute, University of Auckland, NZ). Key collaborators in influenza A analysis: O. Pybus (University of Oxford) and Yi Guan (University of Hong Kong). C Fraser and N Ferguson (Imperial College, Faculty of Medicine) led the 2009 paper [5] to which Rambaut contributed BEAST expertise.

3. References to the research

1. Drummond AJ & Rambaut A (2007) BEAST: Bayesian Evolutionary Analysis Sampling Trees. *BMC Evolutionary Biology* **7**, 214. doi:10.1186/1471-2148-7-214
   3625 Scopus citations at 2nd October 2013
2. Rambaut A, Pybus OG, Nelson MI, Viboud C, Taubenberger JK & Holmes EC (2008) The Genomic and Epidemiological Dynamics of Human Influenza A Virus. *Nature* **453**, 615-619. doi:10.1038/nature06945. **296 Scopus citations at 2nd October 2013**
3. de Oliveira T, Pybus OG, Rambaut A, Salemi M, Cassol S, Ciccozzi M, Rezza G, Gattinara GC, D'Arrigo R, Amicosante M, Perrin L, Colizzi V & Perno CF. (2006) Molecular Epidemiology: HIV-1 and HCV Sequences from Libyan outbreak. *Nature* **444**, 836-837. [equal first authorship]. doi:10.1038/444836a. **49 Scopus citations at 2nd October 2013**
4. Smith GJD, Vijaykrishna D, Bahl J, Lycett SJ, Worobey M, Pybus OG, Ma SK, Cheung CL, Raghwani J, Bhatt S, Peiris JSM, Guan Y & Rambaut A (2009) Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic. *Nature* **459**, 1122-1125. doi:10.1038/nature08182. **714 Scopus citations at 2nd October 2013**
5. Fraser C, Donnelly CA, Cauchemez S, Hanage WP, Van Kerkhove MD, Hollingsworth TD, Griffin J, Baggaley RF, Jenkins HE, Lyons EJ, Jombart T, Hinsley WR, Grassly NC, Balloux F, Ghani AC, Ferguson NM, Rambaut A, Pybus OG, Lopez-Gatell H, Apluche-Aranda CM, Chapela IB, Zavala EP, Guevara DME, Checchi E, Garcia E, Hugonnet S, Roth C & The WHO Rapid Pandemic Assessment Collaboration. (2009) Pandemic Potential of a Strain of Influenza A (H1N1): Early Findings. *Science* **324**, 1557-1561. DOI: 10.1126/science.1176062
   841 Scopus citations at 2nd October 2013

4. Details of the impact

BEAST software has been widely used in non-academic contexts to provide analysis of the evolutionary and epidemiological dynamics of infectious disease agents such as viruses and bacteria. The impact varies from direct application of BEAST to resolve specific questions of infectious disease dynamics with commercial or other impact, to using the understanding derived from this research to inform large-scale public health policy. Overall, BEAST has some 1500 users including commercial users. This case study highlights impact in a number of different contexts.

Impact on public policy and services: Rambaut led a workshop at the Centres for Disease Control & Prevention (CDC) in 2008, teaching BEAST to virologists. CDC is the US national public health institute that works to protect public health and safety by providing information to enhance health decisions and policy. BEAST is now installed on the CDC central high-performance computing facility and is widely used throughout the CDC for public health assessment [a]. For instance, the Head of surveillance at the HIV/AIDS prevention division of CDC states that:

> "*it has been a crucial tool for our routine work in HIV and related viruses for improving public health". [a]*

The CDC trains its staff in use of BEAST for public health surveillance, risk assessment and response [b]. Training courses which including the application of BEAST are also being provided for health surveillance and disease control organisations on behalf of WHO and the FAO/OIE (Food & Agriculture Organisation/World Organisation for Animal Health) [b].

Influence on public health policy and advisory committees: BEAST has been used to inform public health decision making at early stages of an epidemic or outbreak when little other information is available, most notably in public health assessments and approaches taken by the World Health Organization (WHO) during and after the 2009 human influenza (H1N1 'swine flu') pandemic.

The 2009 Fraser et al. Science paper [5] was published relatively early in the outbreak. The 2011 WHO Report '*Strengthening response to pandemics and other public-health emergencies*' [c] cites the publication of this paper as a key event in their process of assessment for the severity of the outbreak. The paper used BEAST to estimate the time of most recent common ancestor and thus the start date of the outbreak and the basic reproductive number (R_0) for pandemic influenza A(H1N1) from genetic sequence data of the virus. The paper also used other epidemiological analyses and estimates of the initial outbreak, international spread, and viral genetic diversity to make an early assessment of transmissibility and severity. The analysis produced using BEAST provides part of a full preliminary analysis contained in this key paper, and the subsequent impact from the paper includes elements directly derived from BEAST as well as from other aspects of the research. For instance the reproductive number R_0 was explicitly referenced in the WHO document '*Considerations for assessing the severity of an influenza pandemic*' published in the WHO Weekly Epidemiological Record (WER) on 29 May 2009 [d]. This summarized the data available to WHO to date including the modelled data from BEAST, addressed how countries should respond in view of the emerging epidemiological, clinical and virological characteristics of the pandemic virus, and noted that WHO provides such information 'to allow countries to tailor their response measures as needed'. A subsequent issue of WER [e] 'summarizes some of the key observations … that may inform preparations being made for the winter influenza season'; this again references paper [5] and the A(H1N1) reproductive number as a key piece of information relevant to such preparations. Rambaut's 2008 [2] and 2009 [4] Nature papers on human influenza A dynamics (primarily H1N1 and H3N2), and also paper [5], were cited in the 2009 WHO Report on "Acute Respiratory Infections" [f] which discusses WHO policy on influenza vaccines and analyses the nature of the 2009 epidemic (including the virology) and its implications for future policy.

Impact on outcomes in criminal and civil litigation: Analysis using the BEAST software has been used in international court cases. The expert report for a large criminal case in Valencia, Spain where an anaesthetist was convicted of infecting hundreds of patients with hepatitis C virus, used BEAST in determining the likely nature and timing of the infection, which was important evidence in determining responsibility [g]. BEAST analysis was also used in expert witness testimony in a patent dispute brought in Norway for fish virus vaccines between Intervet International and Pharmaq AS, where the use of BEAST demonstrated the genetic separation of two virus strains at the core of the case [h].

Impact on society and culture; enhanced awareness of health issues: Rambaut maintained a public website during the swineflu pandemic of 2009 [http://tree.bio.ed.ac.uk/groups/influenza/] on which he published all available current analyses of the pandemic, including analyses using BEAST (e.g. http://tree.bio.ed.ac.uk/groups/influenza/wiki/8fa90/_Outbreak_molecular_epidemiological_analysis_4_May_2009_-_Andrew_Rambaut.html).

This site was frequently used by journalists [i] and Rambaut was interviewed for or quoted in more than 20 general interest, news and popular science publications including 'Wired' magazine (http://www.wired.com/wiredscience/2009/04/swinefluupdate/), the AAAS News (http://news.sciencemag.org/2009/07/pandemic-h1n1-virus-canadian-pigs-smells-odd) and 'LiveScience' (http://www.livescience.com/3668-swine-flu-evolved-unnoticed-years.html). The 2009 Nature paper [4] was quoted in international news media including the China Post (http://www.chinapost.com.tw/health/infectious-diseases/2009/06/13/212080/H1N1-flu.htm) and United Press International (http://www.upi.com/Science_News/2009/06/11/Swine-flu-development-timescale-analyzed/UPI-42611244736329/). Rambaut maintains an ongoing blog http://epidemic.bio.ed.ac.uk/ with news and data on epidemics including the outputs of BEAST analyses which has a wide public audience; for instance on 3rd April 2013, the site had 3,600 hits in one day after Carl Zimmer (Science writer, journalist, columnist at NY Times) tweeted:

> **@carlzimmer:** *If you like watching science unfold in real time, check out this **#H7N9** flu evolution wiki http://epidemic.bio.ed.ac.uk/influenza_H7N9 Data uploaded ASAP*

5. Sources to corroborate the impact

The Tiny URLs provide a link to archived web content, which can be accessed if the original web content is no longer available

a) A corroborating statement is available from the Head of Retrovirus Surveillance in the Division of HIV/ AIDS Prevention, confirming that CDC uses BEAST for public health improvement.

b) A corroborating statement is available from the expert (Duke/NUS medical school) contracted to provide training in BEAST to CDC and WHO.

c) WHO Report 'Strengthening response to pandemics and other public-health emergencies: Report of the Review Committee on the Functioning of the International Health Regulations (2005) in relation to Pandemic (H1N1) 2009' http://apps.who.int/gb/ebwha/pdf_files/WHA64/A64_10-en.pdf Key event cited in appendix VII (p. 163) or http://tinyurl.com/l45rxkn

d) Considerations for assessing the severity of an influenza pandemic. WHO Weekly Epidemiological Record, 2009, 84(22):197–202. http://www.who.int/wer/2009/wer8422.pdf (especially p. 200,Table 1) or http://tinyurl.com/k46r94u

e) Transmission dynamics and impact of pandemic influenza A (H1N1) 2009 virus. WHO Weekly Epidemiological Record, 2009, 84(46):481–484. http://www.who.int/wer/2009/wer8446.pdf (especially p.483 and Table 2) or http://tinyurl.com/lufugvl

f) WHO Report "Acute Respiratory Infections" September 2009: part 2 http://www.who.int/vaccine_research/diseases/ari/en/index1.html (ref 85); Part 6 or http://tinyurl.com/m5mukje
http://www.who.int/vaccine_research/diseases/ari/en/index5.html (refs 408, 420, 438) or http://tinyurl.com/kz7oss7

g) Valencia HepC infection trial: Expert report (in Spanish) referencing BEAST analysis [available on request].

h) Fish vaccine patent court case: The judgement (with English translation) referencing the BEAST analysis and testimony [available on request]

i) Selected publicly-available news media and magazine articles are cited in the impact section text; other articles are available on request.