# PomBase Community Curation: A Fast Track to Capture Expert Knowledge

Midori A. Harris[1], Kim Rutherford[1], Jürg Bähler[2], Paul Kersey[3], Stephen G. Oliver[4], Valerie Wood[1]

[1]Department of Biochemistry, University of Cambridge, Sanger Building, 80 Tennis Court Road, Cambridge CB2 1GA, UK; [2]Department of Genetics, Evolution and Environment, and UCL Cancer Institute, University College London, London WC1E 6BT, UK; [3]EMBL Outstation —European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton CB10 1SD UK; [4]Cambridge Systems Biology Centre and Department of Biochemistry, University of Cambridge, Sanger Building, 80 Tennis Court Road, Cambridge CB2 1GA, UK

**Abstract:**
PomBase, scheduled for launch in summer 2011, will be a new, comprehensive online resource for researchers working on *Schizosaccharomyces pombe*, combining structural and functional annotation, literature curation and access to large-scale data sets. Contributions from the *S. pombe* research community will be the primary source of curated data in PomBase. Following a successful pilot annotation project in which several laboratories contributed annotations — including GO annotations, phenotypes, genetic and physical interactions, and domain data — based on their recent publications, we are now developing a web-based tool to allow the entire *S. pombe* community to participate in PomBase curation. This tool for community annotation is being piloted by *S. pombe* community, but will also be made available to other communities of researchers working with model organisms. Thus, by participating in this programme, *S. pombe* researchers will not only benefit their own community, but also further the study of fundamental aspects of biology with other model organisms. We present the prototype curation interface, supporting GO annotations. By streamlining community contributions to PomBase and establishing pipelines to collaborating databases, we will rapidly work towards comprehensive, high-quality coverage of the *S. pombe* literature, reflect the best, most up-to-date knowledge available for this key model organism, and make *S. pombe* research globally visible.

## Why Add Community Curation?

### Accumulating Literature
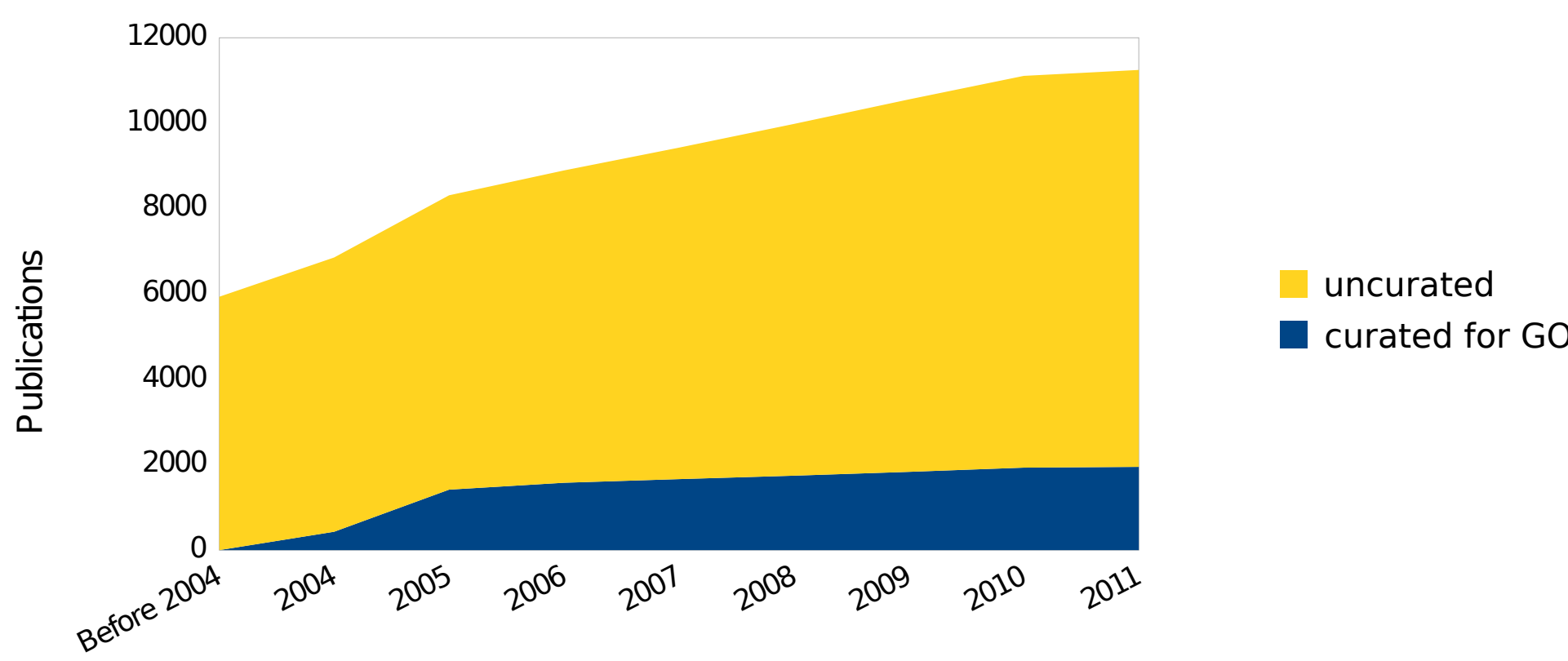


In-depth curation of the entire body of *S. pombe* scientific literature is a primary goal of PomBase, but at present about 7000 publications remain uncurated and the backlog continues to grow.

### Active Community



*S. pombe* is ideal for introducing community curation: researchers contributed annotations based on ~90 papers in a successful pilot project.

## A New Tool for Community Curation

To complement the efforts of PomBase's dedicated database curators, we are developing a web-based curation tool that will allow researchers to contribute annotations directly to PomBase.

### GO annotation in the curation tool



**Tool benefits:**
- easy-to-use interface, accessible on web
- capture expert knowledge direct from community & convert into specific, accurate annotations
- uses controlled vocabularies (ontologies) to ensure consistency
- automated error checking / quality control
- extensible — can add more data types, more controlled vocabularies, etc.
- co-ordinate curation with publication

### Beyond PomBase

#### Data Propagation
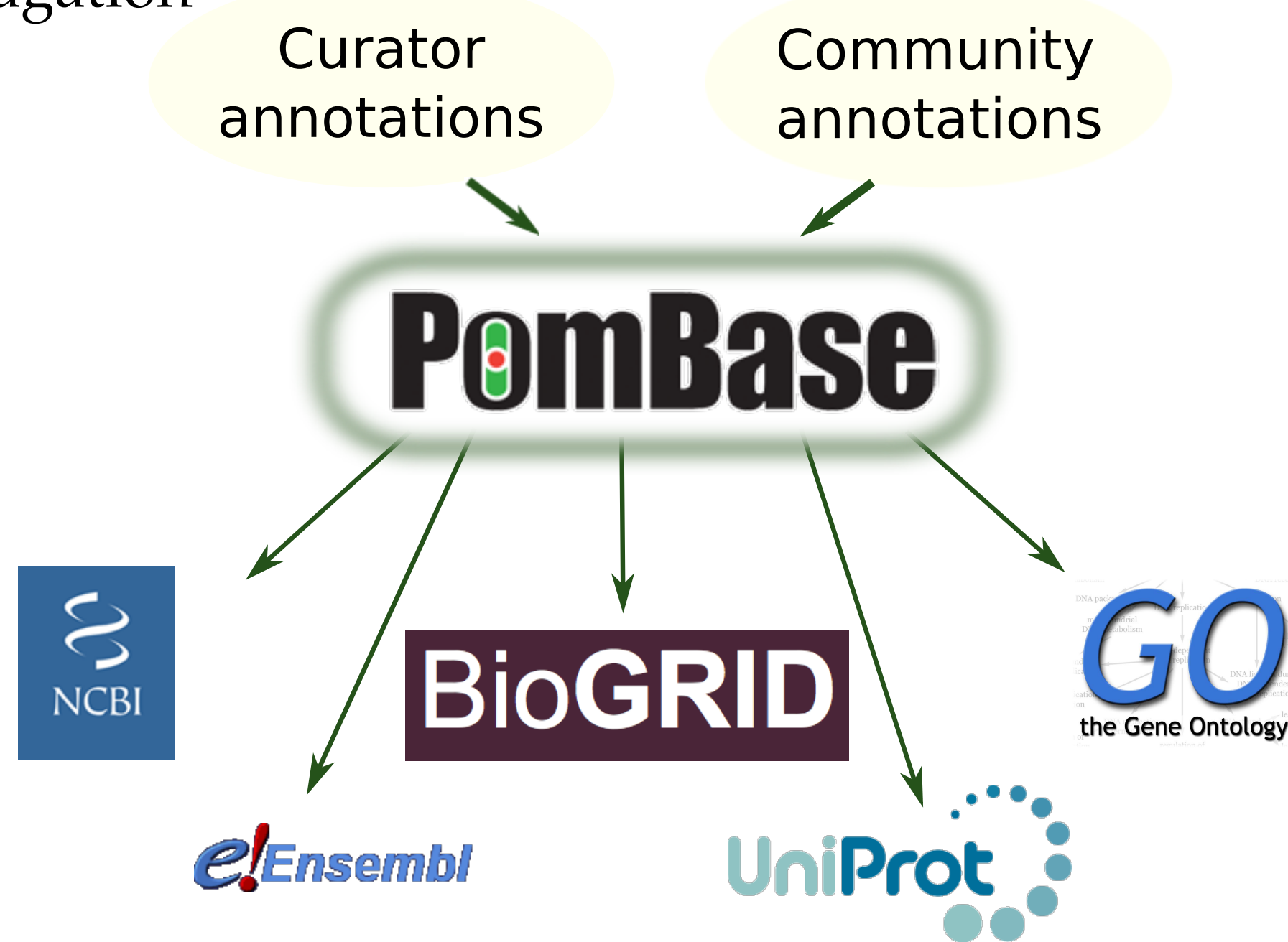


Annotation data go into PomBase via the curation tool, and are then automatically made available to several other databases, ensuring that *S. pombe* research results are visible throughout the online biomedical data sphere.

#### Expanded Community Curation

- More data types will be available for *S. pombe*
  - Phenotypes
  - Modifications
  - Interactions
- The curation tool itself will be made available to researchers working on other model organisms, such as nematodes, fungi and plants.