# WormBase: a comprehensive resource for the genetic model system *C. elegans*

Abigail Cabunoc, Norie de la Cruz, Xiaoqi Shi, Todd W. Harris, Lincoln D. Stein

Ontario Institute for Cancer Research, MaRS Centre, South Tower 101 College Street, Suite 800 Toronto, Ontario, Canada M5G 0A3

C. elegans and cancer

Studies in the model system *Caenorhabditis elegans* have led to the discovery of multiple biological pathways relevant to cancer pathogenesis, including apoptosis, the regulatory roles of micro RNAs, and fundamental cell signaling and differentiation pathways that are dis-regulated in cancer.
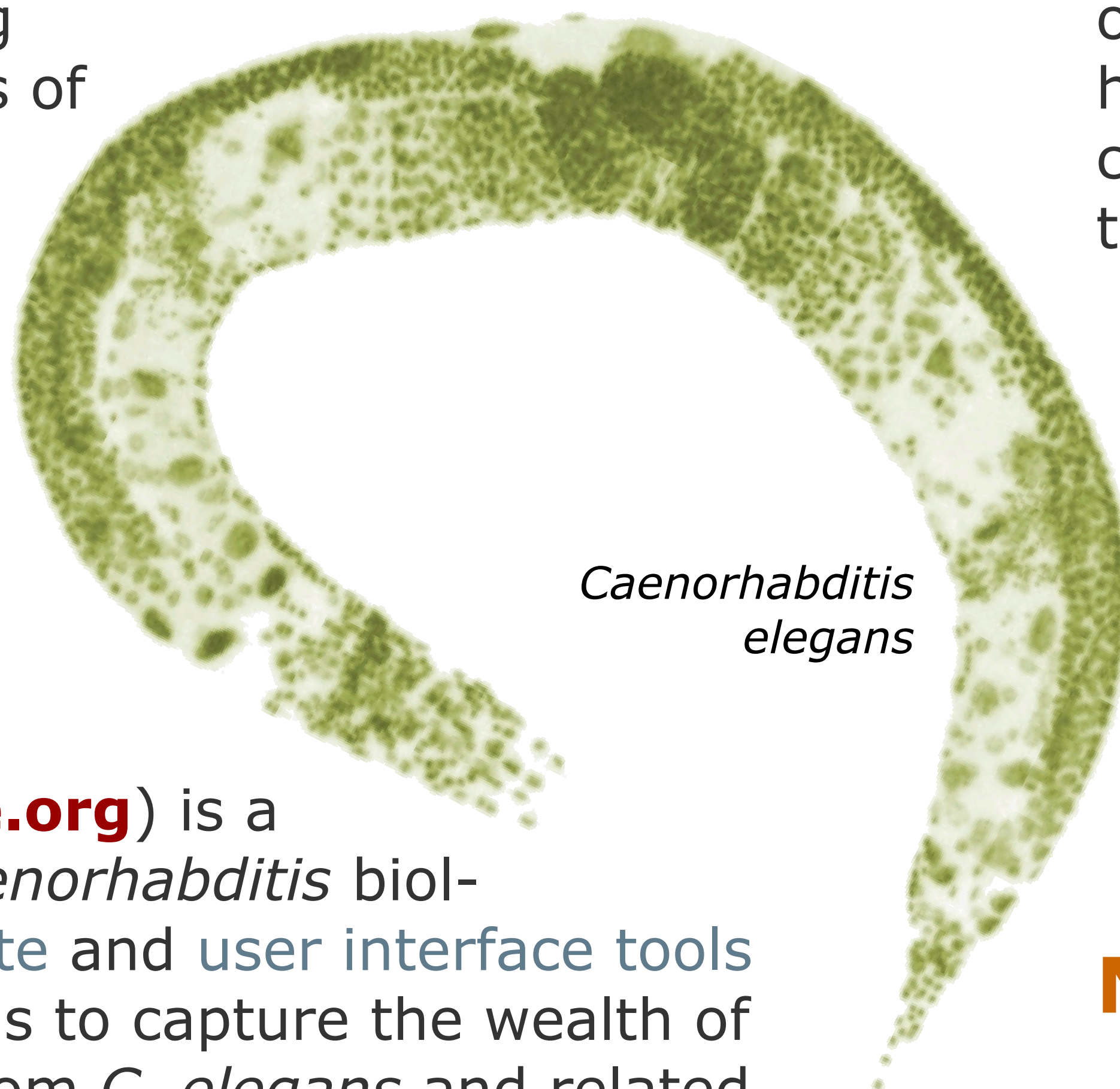
Caenorhabditis elegans

WormBase

WormBase (**www.wormbase.org**) is a central data repository for *Caenorhabditis* biology. OICR manages the website and user interface tools for the project. Our objective is to capture the wealth of experimental data available from *C. elegans* and related nematodes, place it into a rich information discovery space, facilitating new insights into the system.
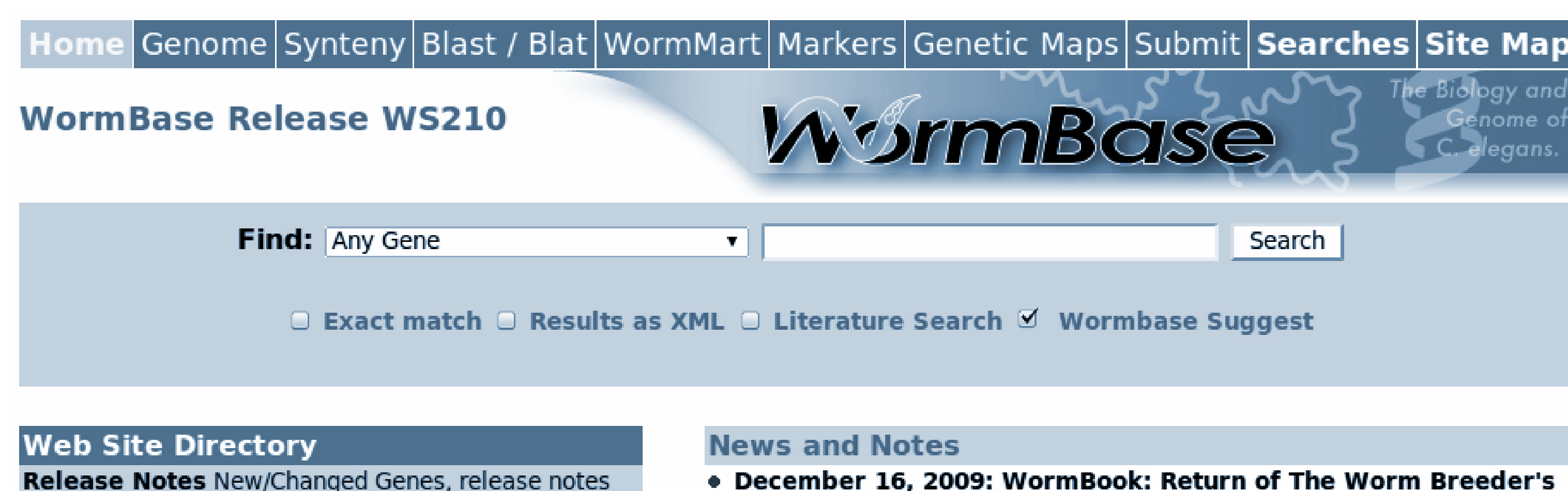
Although initially created as a service to the *C. elegans* research field, WormBase has expanded to include

- complete genomic sequence,
- gene predictions,
- orthology assignments,
- and syntenic alignments

for related species.

Recent additions

Human/nematode ortholog associations that provide easy entry points for those studying human disease processes in model systems

Metabolic pathway information from the Reactome project (www.reactome.org)

Integration of large-scale DNA sequence feature data from the modENCODE project (www.modencode.org).
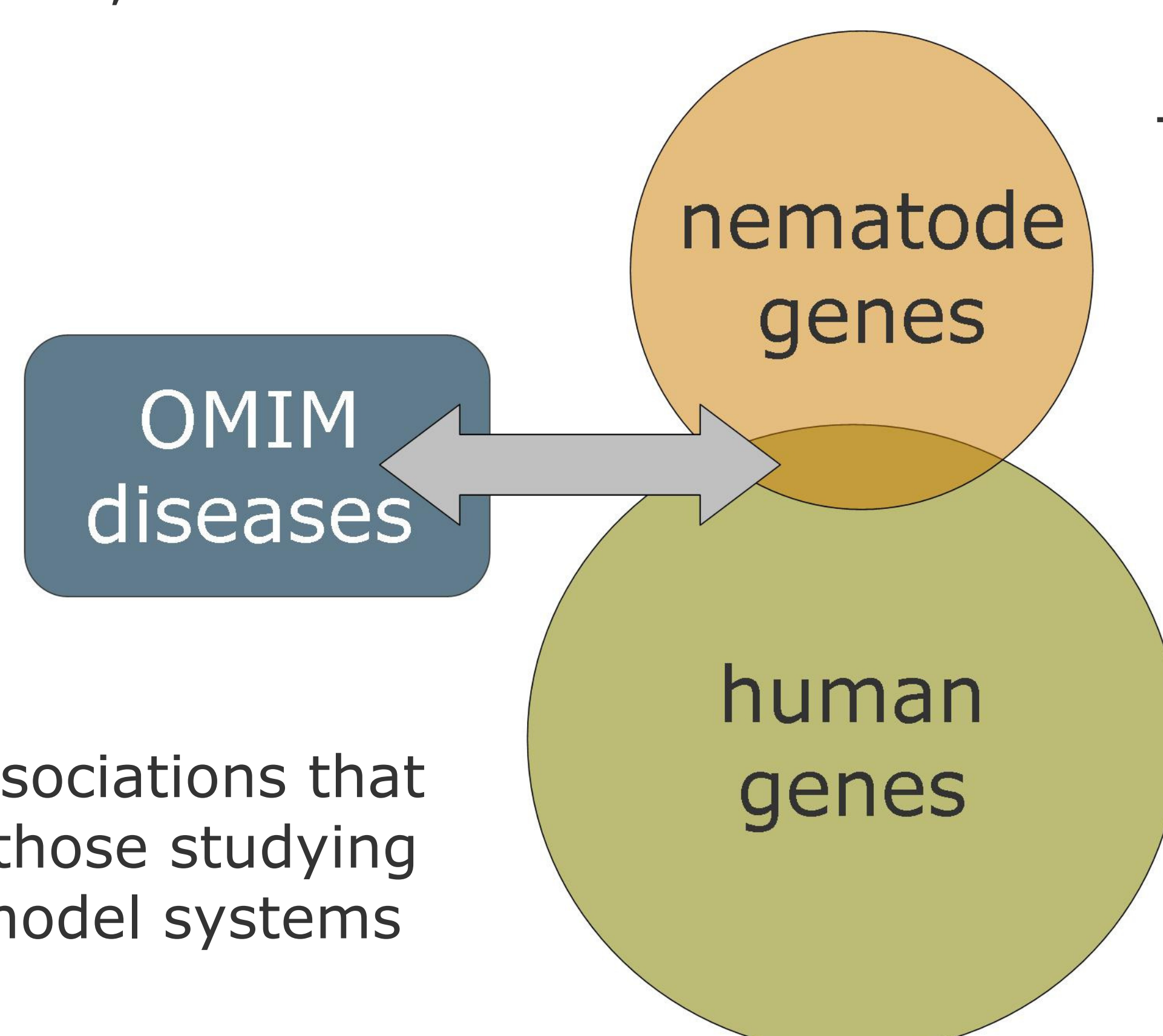
New challenges

The data at WormBase spans great breadth and depth of biology. This constantly growing bolus of data along with an increased user base has revealed areas for improvement in the current architecture. Several aspects of the resource are handling far more data than originally intended, leading to un-wieldy and hard to maintain areas.
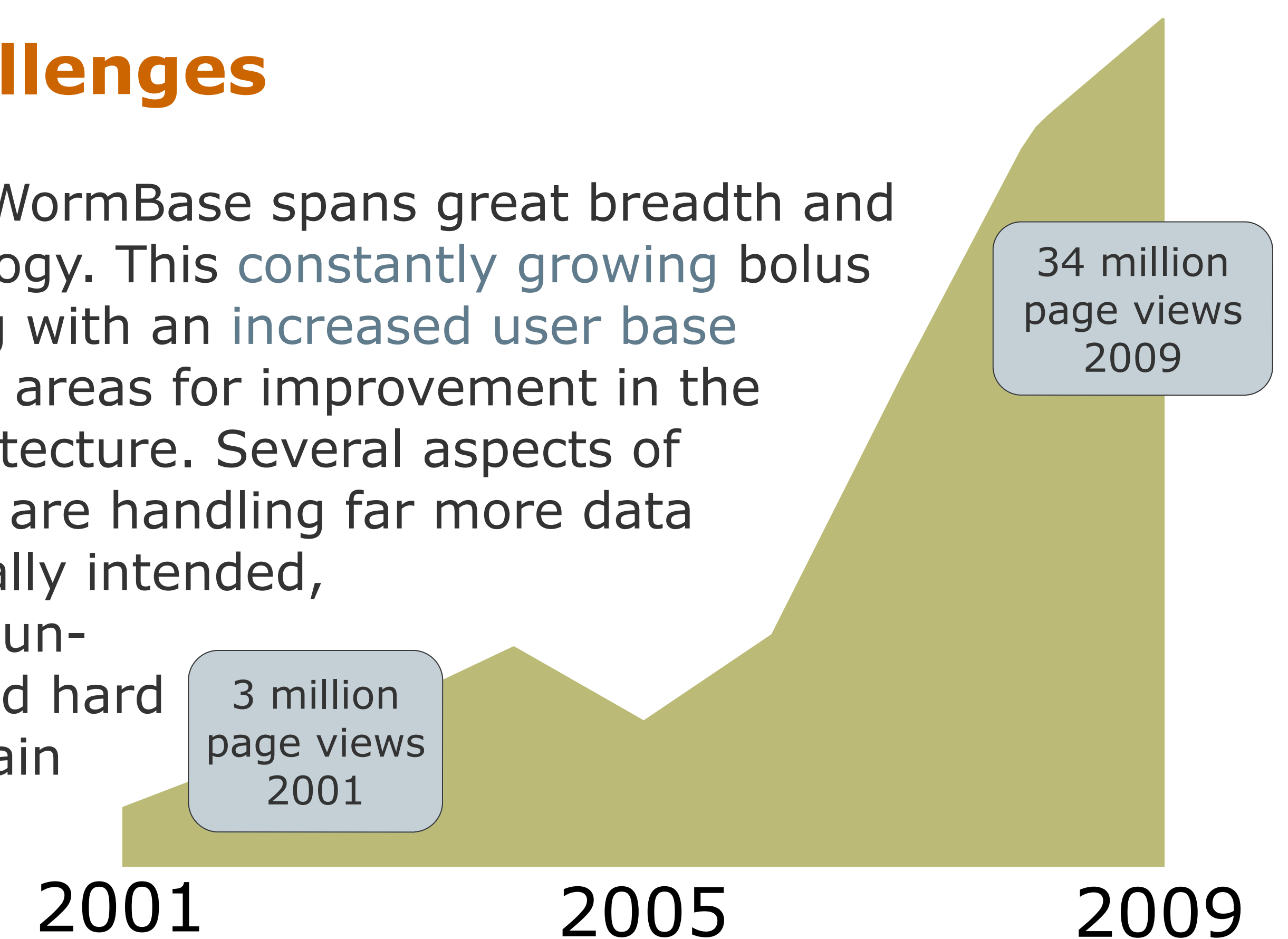
Next generation WormBase

We are refining our system into a modern and modular information architecture. Using the model-view-controller architectural pattern we will isolate different aspects of the application logic, enabling easier testing and maintenance.
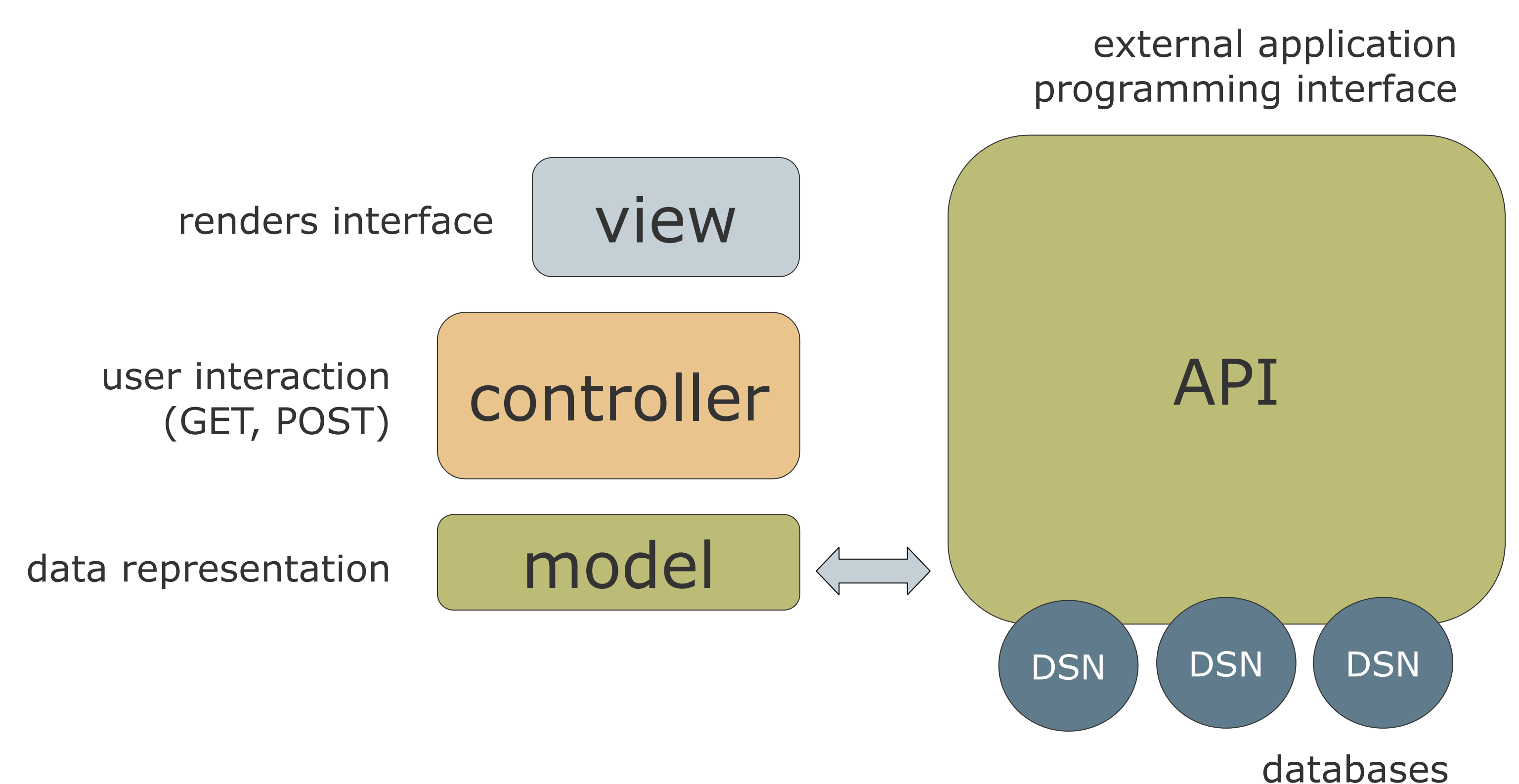
This redesign includes revamping the user interface to make it more flexible and customizable for end users, making it easier to build custom queries of the database, and leveraging cloud data services when appropriate.

modern web application

- fast
- flexible
- easy to maintain
- easy to contribute
- scalable

mononucleosome binding site data from the **modENCODE project** displayed on the **Generic Genome Browser 2.0**