

# The PhenoDB Ecosystem: Open-Source Web-based Tools for Genomic Discovery

Stewart Jester<sup>1,2</sup>, Sean M.L. Griffith<sup>1,2</sup>, Renan Martin<sup>1</sup>, Laura Vail<sup>1,2</sup>, Corina Antonescu<sup>1</sup>, Alice Sanchez<sup>1,2</sup>, Nara Sobreira<sup>1</sup>

<sup>1</sup>Department of Genetic Medicine // <sup>2</sup>Johns Hopkins Genomics  
Johns Hopkins University School of Medicine, Baltimore, MD, USA



## An Introduction to the PhenoDB Ecosystem

PhenoDB is a web-based tool that provides methods to annotate and analyze the variants of patients and their family members with rare Mendelian diseases.

The PhenoDB software ecosystem additionally includes GeneMatcher and VariantMatcher, respectively developed to enable connections between patients, clinicians, and researchers from around the world who share an interest in the same genes and variants.

GeneMatcher and VariantMatcher are part of the Matchmaker Exchange, with the principal goal of sharing genomic and phenotypic data to facilitate the discovery of novel disease variants and genes.

VariantMatcher allows sharing of variant-level and phenotypic data from whole exome and genome sequencing.

These three tools share a development team and a central codebase, allowing researchers to use them in conjunction.

## PhenoDB 2.0: Robust and Portable

Web	Django 5 and Python 3	The PhenoDB ecosystem is kept up to date with the latest version of Django, ensuring security and portability, with additional code written in Python 3.
Structured Data	MariaDB 10	Structured data that forms the back-end for the web tools are stored in MariaDB, with MySQL providing an additional option. Data access is strictly controlled.
Searchable Data	SOLR 9	Data is searchable across many strata, such as variant and submission, allowing the user to more easily discern which genes are candidates for further investigation.
Application Portability	Docker Compose	Containerization of the application layers of the PhenoDB ecosystem allows for rapid deployment of new versions and simplified installation via Docker Compose.
Platform	Azure or Linux	The containerized applications of the PhenoDB ecosystem are tested on both Azure and Linux, though containerization allows a broad spectrum of options.

## New Analysis Capabilities and Integrations

### VariantMatcher

VariantMatcher is now connected to MyGene2/Geno2MP, and Franklin, with the development team regularly working with The GA4GH Beacon Work Stream on API features and new connections.

### GeneMatcher

As of October 2024, GeneMatcher has been cited in 909 publications describing 689 (76%) novel disease gene discoveries.

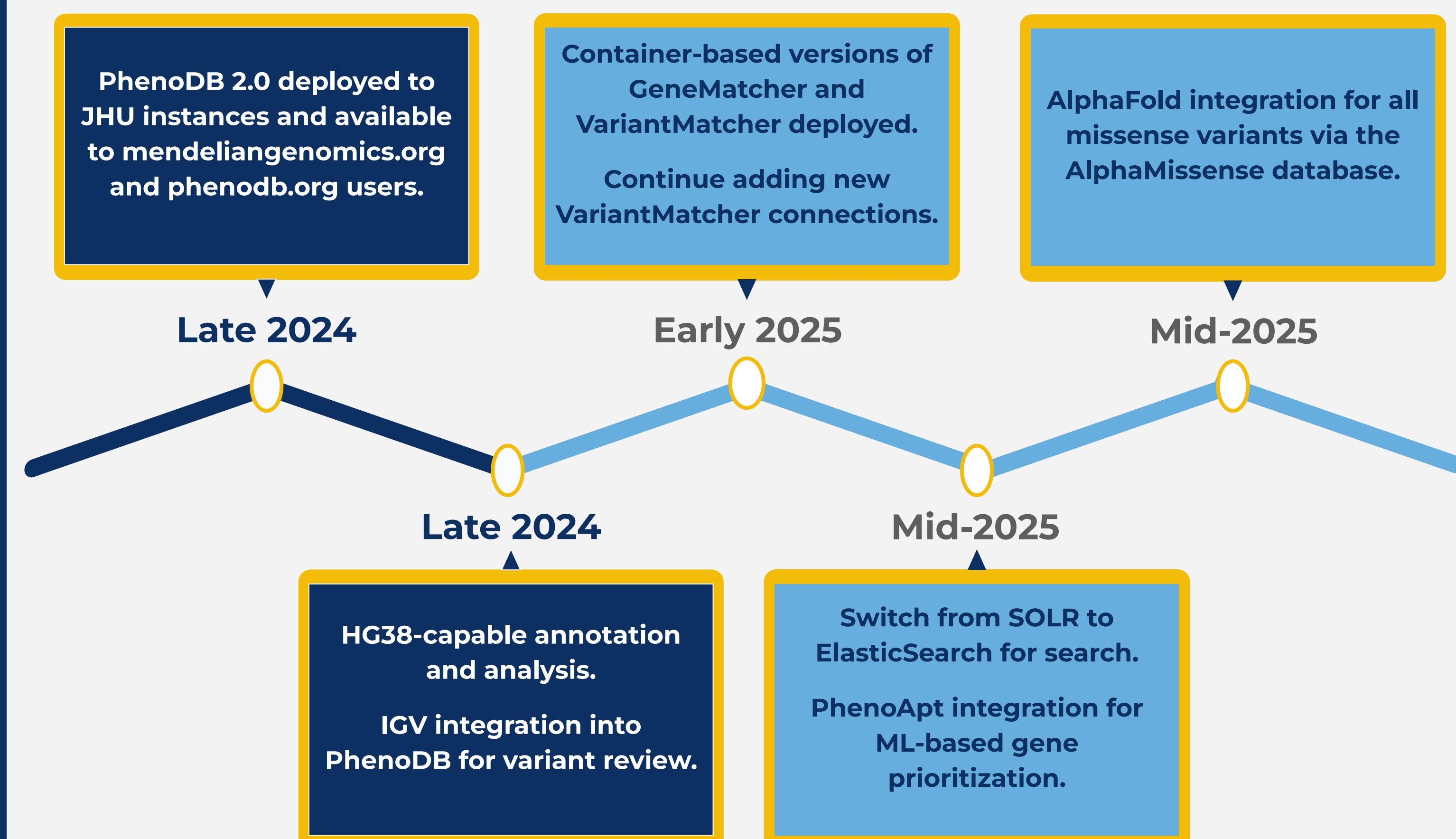
### CAVATICA Integration

Annotation and analysis features from PhenoDB are now available as apps for any CAVATICA user. Analysis parameters include MAF cutoff, RefGene gene location, and multiple inheritance types.

### Analysis and Annotation

PhenoDB 2.0 has significantly enhanced annotation and analysis capabilities for variant filtering and prioritization. In the analyze design tab, users can prioritize variants based on the gene-disease association in OMIM, gnomAD constraints scores (missense Z-score and pLI score), CADD, and RAVEL scores.

## Next Steps and New Directions



## PhenoDB Pilot Installation Program

### The Benefits of Collaboration

With a greatly simplified installation process, PhenoDB 2.0 is accessible to research and clinical genome centers around the world. Over the past six months, we have engaged in a pilot installation program with a group of interested centers to install PhenoDB 2.0. This has given these centers access to the annotation and analysis tools of PhenoDB and simultaneously allowed the PhenoDB development team to fine tune our installation process, while creating solutions for the range of challenges that other teams may face.

Our ultimate goal is to make our tools available to as many groups as possible, and we cannot do that without close work with partner organizations. We attempt to directly support any interested organization, allowing for a wide variety of security requirements, technical skills, and resource availability. We also attempt to support groups that need to adapt the PhenoDB installation process to match their required technology stack, such as using Singularity instead of Docker or AWS instead of Azure.

### Do you want to install PhenoDB?

#### What do you need?

We recommend either access to an Azure account **or** a dedicated Linux system with at least 150 GB of storage and 32 GB RAM, with MariaDB or MySQL, and Docker installations. We are also glad to work with you to try other cloud providers or operating systems.

#### Who should you contact?

Nara Sobreira at nsobrei2@jh.edu  
Sean M.L. Griffith at sgriff21@jh.edu

#### What should you expect?

You will be invited to the monthly open PhenoDB installation session, where we will introduce ourselves, discuss your needs, how you'd like to install PhenoDB, and the technical expertise available to you.

We will then make the PhenoDB installation package and codebase available to your team and work with you directly to address any installation issues, validate full installation, and help users learn how to use PhenoDB.

