



# Biosciences Working Group

Jason Haga

AIST

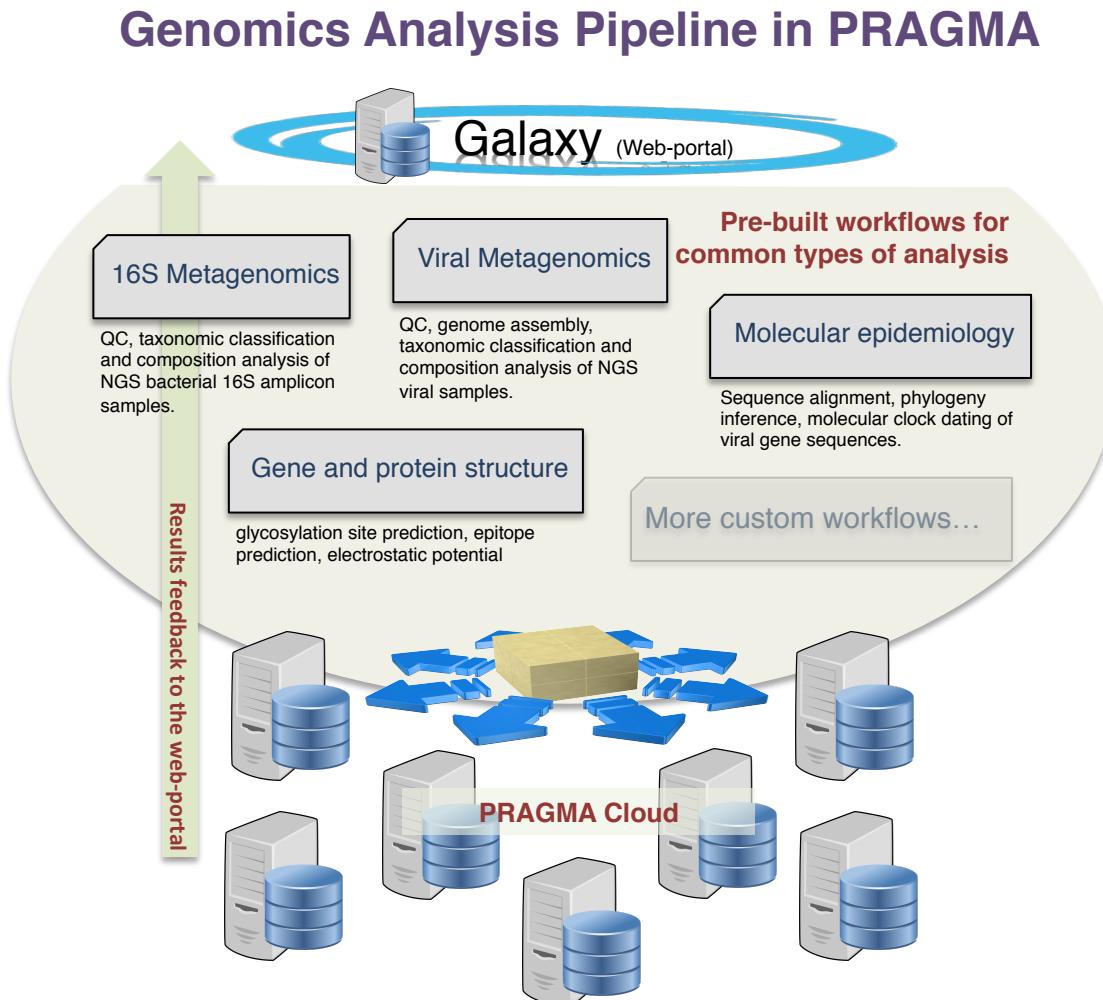
PRAGMA 30, Jan 28, 2016

# Bioscience Activities

- Genomics Analysis Pipeline
  - Creating resources to enable genomics analysis
- Natural Products Discovery Portal
  - Create a unified platform for drug discovery
- Both driving applications of PRAGMA Cloud
  - Cross-working group interaction

# Genomics Analysis Pipeline

Infectious disease as domain



Phase 1: Set up web-portal and computing infrastructure for Galaxy and genomics tools in PRAGMA resources.

Phase 2: Build workflows for common types of genomic analysis, initially for field of infectious disease research and then extend to other fields with the contributions of other PRAGMA participants.

Phase 3: Testing using case studies contributed by PRAGMA participants.

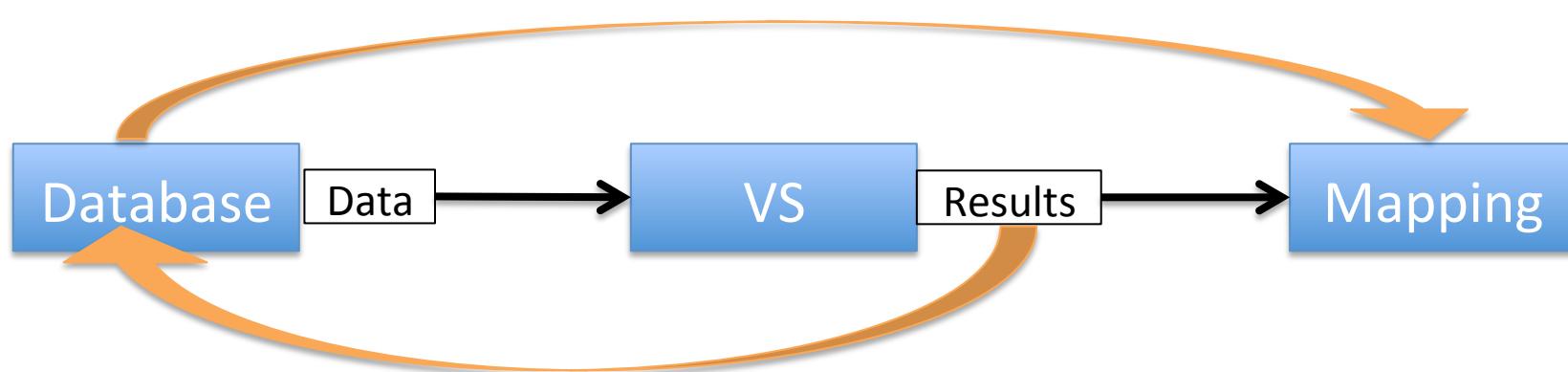
Phase 4: Education via training workshops and invite wider contributions to building common analysis workflows and application with PRAGMA participants' data.

# Since PRAGMA29

- Genomics Analysis Pipeline
  - Created KVM-based VM with Biolinax (Kohei)
  - Preliminary testing of VM (Tommy)
  - VM is now on UCSD resources and in scheduler (Shava)
  - Starting to build up PRAGMA Cloud resources (AIST)

# PRAGMA Natural Products Discovery Platform

- Components
  1. Natural products database
  2. Virtual screening
  3. Biodiversity mapping
- What are the infrastructure requirements for each component?
  1. Storage, data format
  2. Network, compute nodes
  3. Mapping tools, visualization



**BASIS DATA TANAMAN OBAT INDONESIA**

<http://herbaldb.farmasi.ui.ac.id/v3/index.php>

Beranda Daftar Spesies Daftar Senyawa FAQs Tentang Kami IND ENG

Database Senyawa Aktif Tanaman Obat Indonesia

**Box of Search**

Kategori Pencarian  
Kunci Pencarian  
Search

**Members Login**

Ingin Mendaftar? Login

**Avocado (Alpukat)**  
Tumbuhan Avocado berasal dari Meksiko dan Amerika Tengah dan kini banyak dibudidayakan di Amerika Selatan dan Amerika Tengah sebagai tanaman perkebunan monokultur dan sebagai tanaman pekarangan di daerah-daerah tropika lainnya di dunia.

3 / 3

Stu

# HerbalDB (Indonesia)

[http://kanaya.naist.jp/KNAPSAck\\_Family/](http://kanaya.naist.jp/KNAPSAck_Family/)

**“KNAPSAck” Family**

Since 2008.07

**NAIST**

**KNAPSAck Metabolomics**

**3D** Since 2012.11    **Core System** Since 2004.04    **Search Engine** Since 2008.12

**Pocket** Search for Functional Species

<b>Food &amp; Health</b>	<b>Crude Drug</b>	<b>Biology</b>
Lunch Box 食用データベース Since 2008.07	WorldMap 世界の薬用植物 データベース Since 2009.06	Biological Activity Natural Activity Since 2011.06
DietNavi 病気予防データベース Since 2012.11	KAMPO 漢方薬 生薬 データベース Since 2008.09	Biological Activity Metabolite Activity Since 2013.01
FoodProcessor 加工食品データベース Since 2012.11	JAMU IndonesiaHerb データベース Since 2009.11	
DietDish 食べ合わせデータベース Since 2012.11	Tea Pot ハーブ データベース Since 2011.06	
MARCHÉ 毎データベース Since 2014.04	Skewered KNAPSAck 串刺し検索 Since 2010.10	

**Picnic** Gene Annotation    **Strap** Correlation Coefficient

Arabidopsis Since 2008.04    Arabidopsis Since 2009.09

Bacillus Since 2008.05    Bacillus

Human Since 2009.03    Human

**Pickaxe** Metalloprotein Database    **Motorcycle** Metabolic Pathway

MetalMine Since 2009.09    代謝データベース Since 2011.08

**Bicycle** Algae Metabolic Pathway    **Bicycle** 代謝データベース Since 2013.09

Instruction Manual(Japanese)

Instruction Manual(English)

# KNAPSAck (Japan)

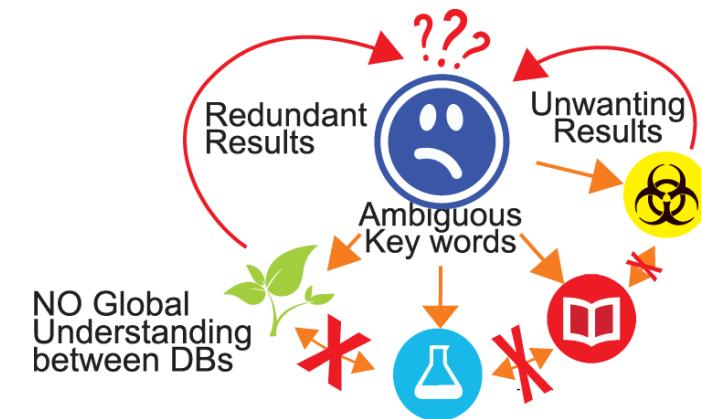
# Since PRAGMA29

- Natural Products Discovery Portal
  - Database refinement/cleanup (Arry)
  - VietHerb database (Ly)
  - Autodock in Docker containers (Heru)
  - DOCK in Docker containers (Weicheng, Max)

# Ontology-based Herb Database System for Pharmaceutical Science and Drug Development

## Background and Challenge

- Huge but fragmented DBs for herbs, metabolites, and diseases
  - ➡ Big barrier for interdisciplinary studies
- Widely used DBs not semantically related
  - ➡ No knowledge sharing for analogous terminologies
- Terminologies ill-defined
  - ➡ Unwanted/redundant/incorrect results



## Objectives and Contributions: VietHerb

- Integration of fragmented resources to unify and correlate Knowledge of Herbs, metabolites and diseases
- User friendly interface to assist both non-experts and interdisciplinary scientists for best results
- Updating mechanism for anonymous experts to collaboratively contribute to provide new knowledge

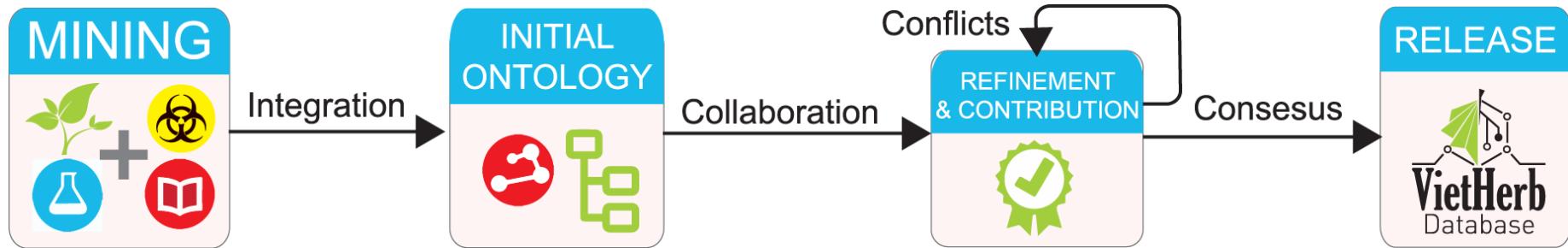
The screenshot shows the homepage of the VietHerb Database. It features a background image of various herbs and a central logo with the text "VietHerb Database". A banner on the right side reads "Most updates on Vietnamese Medical Herbal Database." and includes a link "[Find out more >>](#)". Below the banner is a navigation bar with links: HOME, DICTIONARY, SUGGESTIONS, FAQ, ABOUT, PARTNER, TIẾNG VIỆT. A search bar is located at the top right. The main content area displays a table with the following data:

ID	SCIENTIFIC NAME	Vn-NAME	Eng-NAME	METABOLITES	THERAPEUTIC EFFECTS
HID_00001	Aconitum napellus L.	Áu ô đầu	Monk's-hood [2] aconite wolfsbane, fuzi, monk's blood	Balanine G; T2-epi-Napelline; Songasmine	Animocleptic; Analgesic; Anti-epileptic; anti-convulsive; Anti-convulsive; Anti-shock; Anti-inflammatory; Anti-bacteria; Anti-funga;

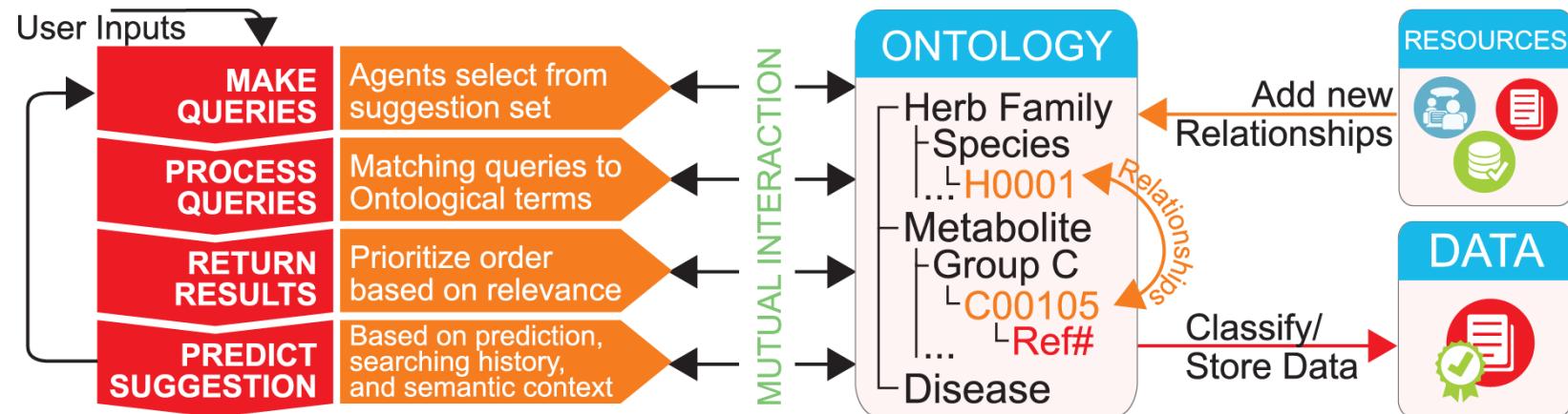
# Ontology-based Herb Database System for Pharmaceutical Science and Drug Development

## Key technologies

- Integration of different sources to build a sound ontology



- Use of ontology for semantic search and information exchange



PI: Ly Le (International Univ., Vietnam National University), FY15:30K, FY15:30K)