COMPARE

Comprehensive Protein Allergen Resource

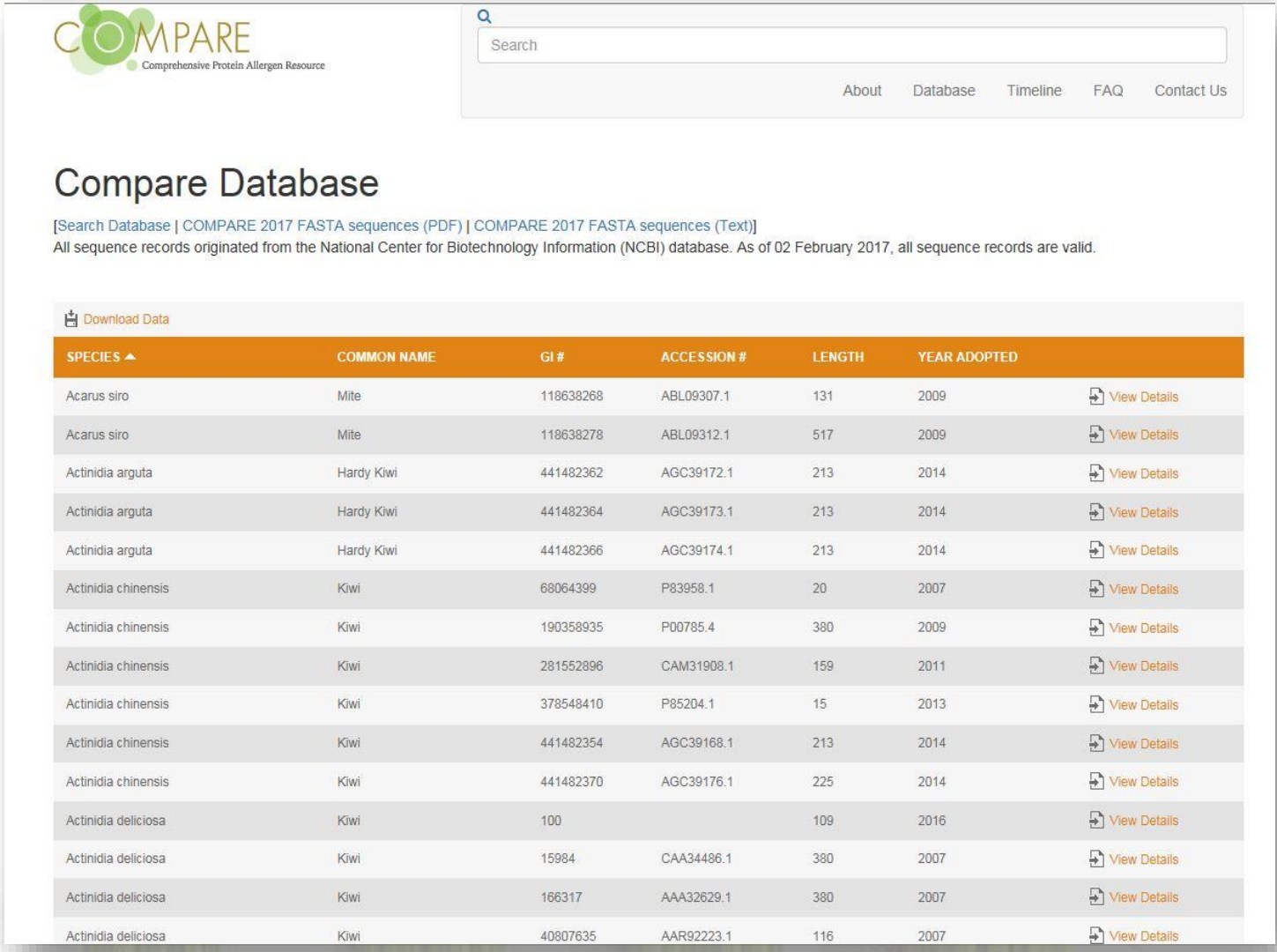
Creating a public resource for protein allergenicity assessment: The COMPARE Database.

Health and Environmental Sciences Institute (HESI)
Protein Allergenicity Technical Committee (PATC)

HESI COMPARE Allergen Database: www.comparedatabase.org

COMPARE
Comprehensive Protein Allergen Resource

- A comprehensive and up to date collection of peer reviewed allergen protein sequences. Builds on a > 10 year database of peer review science.
- Includes known allergens from 2016 and previous years (listed in existing public database AllergenOnline V16) and *new* allergens identified through the COMPARE process.
- Available since February 03, 2017 at www.comparedatabase.org (website hosted and maintained by HESI)
- Updated annually



SPECIES	COMMON NAME	GI #	ACCESSION #	LENGTH	YEAR ADOPTED	
Acarus siro	Mite	118638268	ABL09307.1	131	2009	View Details
Acarus siro	Mite	118638278	ABL09312.1	517	2009	View Details
Actinidia arguta	Hardy Kiwi	441482362	AGC39172.1	213	2014	View Details
Actinidia arguta	Hardy Kiwi	441482364	AGC39173.1	213	2014	View Details
Actinidia arguta	Hardy Kiwi	441482366	AGC39174.1	213	2014	View Details
Actinidia chinensis	Kiwi	68064399	P83958.1	20	2007	View Details
Actinidia chinensis	Kiwi	190358935	P00785.4	380	2009	View Details
Actinidia chinensis	Kiwi	281552896	CAM31908.1	159	2011	View Details
Actinidia chinensis	Kiwi	378548410	P85204.1	15	2013	View Details
Actinidia chinensis	Kiwi	441482354	AGC39168.1	213	2014	View Details
Actinidia chinensis	Kiwi	441482370	AGC39176.1	225	2014	View Details
Actinidia deliciosa	Kiwi	100		109	2016	View Details
Actinidia deliciosa	Kiwi	15984	CAA34486.1	380	2007	View Details
Actinidia deliciosa	Kiwi	166317	AAA32629.1	380	2007	View Details
Actinidia deliciosa	Kiwi	40807635	AAR92223.1	116	2007	View Details

Who will use the COMPARE Database?

Translate

Who?

- **Product developers** providing safety information on novel proteins.
- **Regulatory agencies** responsible for food and feed safety assessments.
- **Clinicians**, medical personnel and caregivers in the allergy field.
- **The Public**, who may be interested in identifying sources of allergens.

How?

- Bioinformatics tools such as the FASTA algorithm are utilized to evaluate the degree of similarity between introduced proteins into a crop and known allergens – COMPARE allows these bioinformatic comparisons.
- Researching types of allergens, sequence information and literature associated.
- Staying updated with new allergens identified each year (all users).

Access to a transparent and consensus-based allergen database is a key aspect for supporting public safety.

RATIONALE FOR AN ALLERGEN DATABASE AT HESI

1) Key components of the HESI public-private partnership.

HESI PATC COMPARE

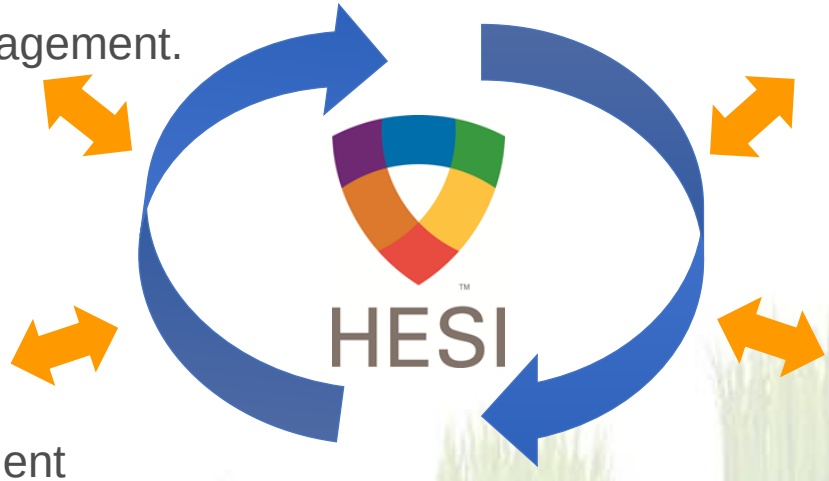
P-P Science Advisory Team

✓ Multiple stakeholders engagement.

Independent Peer Review Panel (PRP)

Allergy Science Expertise

✓ Decisions based on peer-review



Bioinformatics services

✓ Custom algorithm development

JIFSAN
Joint Institute for Food Safety and Applied Nutrition

✓ Rigorous process for data management and tracking (custom information management tool).

✓ Literature procurement

✓ **Independent coordinating body:**

- Scientific program management
- Budget and contracts management with external professional services.

2) Modular process allows multiple sourcing of allergen candidates.

Goal: enhance confidence (no allergens left behind)

3) Transparency & detailed documentation.

What has been done:

- COMPARE overview brochure: at initiation of the project (Q1 2016).
- Detailed process development described in the COMPARE website: within 6 months of the database public release (COMPARE' 2017). http://comparedatabase.org/process-development/
- Announcement letter sent to stakeholders and published on JIFSAN's website.

In preparation:

- Peer-review publication

COMPARE
Comprehensive Protein Allergen Resource

COMPREHENSIVE PROTEIN ALLERGEN RESOURCE

Program Overview – February 2016

ABOUT HESI

HESI, the Health and Environmental Sciences Institute, is a non-profit, global scientific organization dedicated to developing science for a safer, more sustainable world.

HESI's mission is to collaboratively identify and help to resolve global health and environmental challenges through engagement of scientists from academia, government, industry, NGOs, and other strategic partners.

HESI

OVERVIEW

The COMPARE Database, to be made publicly available in January 2017, will be a transparent resource for identification of protein sequences that are known or putative allergens. The COMPARE database will meet needs for allergy safety assessment via an annual updating process that 1) captures new listings of allergen sequences, 2) filters out non-allergen sequences, 3) identifies published literature linked to the identified potential allergen sequences, and 4) verifies that the newly identified sequences have clinical, published evidence of allergenicity based on standardized criteria.

PROGRAM STRUCTURE

The COMPARE database is a collaborative HESI program that combines programmatic support from the Joint Institute for Food Safety and Nutrition – www.jifsan.org at the University of Maryland - with financial support from private sector partners. Additional in-kind support from multiple academic and government partners is engaged throughout the process. A public-private HESI collaborative team will publicly document all database design and search algorithm decisions. An exclusively public-sector expert panel will review the output from a high-throughput bioinformatic pipeline against peer-reviewed publications. The expert panel will then select sequences for inclusion in the database and document their rationale for inclusion.

TIMELINE

The COMPARE database will be released to the public by January, 2017. To enable this timing, a steering team will convene in early 2016 to ensure timelines, process, and resources are aligned. Sequence searching will be conducted in second quarter 2016 and peer-review in third quarter 2016 before final assembly and release.

Conclusion: overall program design and components enable long term, high quality and stable support.

COMPARE
Comprehensive Protein Allergen Resource

Advisory Team

- 20 Public/Private stakeholders.
- General program oversight, QC, user feedback.
- Expert knowledge
- No influence on decisions regarding sequence inclusion or exclusion.

Bioinformatics and Data Communication

- **AccuraSciences**: keyword filtering and algorithm design; sources information for database.
- **University of Maryland / JIFSAN:**
  - Literature search and procurement.
  - Data communication.
  - Review Tool.

JIFSAN
Joint Institute for Food Safety and Applied Nutrition

Program Coordination

- Overall program management.
- Coordination between partners, PRP assembly and management.
- Database 'home' and maintenance.

HESI

Expert Peer Review Panel

- 5 invited public sector allergy experts (from the Universities of Vienna, Copenhagen, and Amsterdam; Mount Sinai, NY; and UC Davis, CA).
- Define inclusion /exclusion criteria.
- Review of sequences and references / publications.
- Provide final decisions on new sequences to include in database.

Participating Organizations

Public Sector:
- Universities of Amsterdam
- University of Maryland / JIFSAN
- U.S. EPA
- U.S. FDA

Private Sector:
- BASF
- Bayer CropScience
- Dow
- DuPont Pioneer
- KWS
- Monsanto
- Syngenta
- Vilmorin & Cie

For more information:

Visit:

✓ HESI PATC website: committee fact sheet, events and publications:

http://hesiglobal.org/protein-allergenicity-technical-committee-patc/

✓ COMPARE Database: www.comparedatabase.org

Contact:

Lucilia Mouriès, HESI PATC Scientific Program Manager:

lmouries@hesiglobal.org