

# BAR ThaleMine

Asher Pasha<sup>1</sup>    Nicholas Provart<sup>1</sup>

<sup>1</sup>Bio-Analytic Resource (BAR)  
Centre for the Analysis of Genome Evolution and Function  
University of Toronto  
Toronto, Canada

January 28, 2020

# BAR: Bio-Analytic Resource

## About us:

- ▶ One of the largest Plant Biology Resources
- ▶ Paper published in 2005<sup>1</sup>
- ▶ Based in University of Toronto, Canada
- ▶ Free to access: <http://bar.utoronto.ca>

## Resources Include:

- ▶ Web applications, web services, databases
- ▶ Gene Expression data ([eFP](#), [ePlant](#))
- ▶ Protein-Protein, Protein-DNA Interactions ([AIV2](#))
- ▶ Protein structures ([ePlant](#))

---

<sup>1</sup>Toufighi et al. 2005.

# Araport: Arabidopsis Information Portal

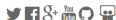
A centralized resource for Arabidopsis data<sup>2</sup>

- ▶ Started in 2013 under the leadership of Prof. Chris Town (JCVI)
- ▶ JCVI, TACC, University of Cambridge
- ▶ Drupal 7 based [araport.org](http://araport.org) page
- ▶ ThaleMine (InterMine instance)
- ▶ App store on araport.org
- ▶ [JBrowse](#)
- ▶ Araport11 re-annotation project
- ▶ Single Sign on to all Araport resources
- ▶ RESTful API for scientific data
- ▶ Community outreach (workshops, conferences, etc)

---

<sup>2</sup>Krishnakumar et al. 2015.

# Araport: Arabidopsis Information Portal



**ARAPORT**  
ARABIDOPSIS INFORMATION PORTAL

[About ▾](#)[Data ▾](#)[Tools ▾](#)[Community ▾](#)[Developer ▾](#)[Help ▾](#)[Log in ▾](#)[ALL](#)[THALEMINE](#)[JBROWSE](#)

Search ThaleMine, JBrowse and other site content

☒ Search by identifiers or keywords e.g. AT4G19020 or epigenetic

## New to Araport?

Araport is a one-stop-shop for *Arabidopsis thaliana* genomics. Araport offers gene and protein reports with orthology, expression, interactions and the latest annotation, plus analysis tools, community apps, and web services. Araport is 100% free and open-source. Registered members can save their analysis, publish science apps, and post announcements. [READ MORE](#)

[REGISTER](#)[LOGIN](#)

## THALEMINE



Click to analyze gene lists or drill down to gene reports.

## JBROWSE



Click to browse gene structures and other annotation.

## APP STORE



Click to build your own analysis tools and share your data.

## BAR ePlant



Visual analytic tool for exploring multiple levels of Arabidopsis data.

## Araport Highlights



Official Release Jun 2016

529092123/https://www.araport.org/ Re-annotation



Upcoming Events

ICAR, June 19-23

## News and Job Postings

★ ABRC Funding Update

2016-04-25

## Conferences and Meetings

Frontiers and Techniques In Plant Science

2017-06-30

# Araport: Arabidopsis Information Portal

## BAR contibution to Araport

- ▶ Provided gene expression database used by BAR eFP Arabidopsis Browser
- ▶ Attended Araport workshop/hackathon in November 2014
- ▶ Interactions Viewer - the first community app published on Araport.org
- ▶ Contributed serveral apps and APIs over the next few years.
- ▶ Biggest contribution: BAR ePlant app.

# Araport: Arabidopsis Information Portal

## Funding status

- ▶ Lost all funding in 2016
- ▶ Lost almost all staff
- ▶ New application for funding failed

## Future of Araport team

- ▶ JCVI, TACC, BAR, TAIR, NCGR
- ▶ First meeting at JCVI in March, 2019
- ▶ BAR got ThaleMine, TAIR got JBrowse, NCGR contributed GCV
- ▶ Every thing else was deprecated
- ▶ News Update: [araport.org](http://araport.org)

# Araport ThaleMine to BAR ThaleMine

Araport ThaleMine: A few minor problems

- ▶ Not maintained since 2017
- ▶ Based on InterMine 1.8.5
- ▶ Used modified data
- ▶ Used several custom loader and displayers
- ▶ Some InterMine core and biosource changes where never made upstreams
- ▶ Too difficult to upgrade to InterMine 3.1.2

# BAR ThaleMine: Principles

No initial funding. So we follow these rules

**Simplicity** Keep code to minimum possible lines and contribute as many changes to InterMine (upstream) as possible.

**Least Surprise** Test and upgrade every minor release (Principle of least astonishment)

**Clarity** No modification to data

**Extensibility** Build for the future (Java 11, etc)

**Parsimony** Write loaders and displays only when we have to.



# BAR ThaleMine

- ▶ We started from scratch (using make mine script)
- ▶ Based on Humanmine, FlyMine, and Araport ThaleMine
- ▶ Got most of the loaders working
- ▶ Pushed changed to upstream InterMine
- ▶ So far: ThaleMine v4.1.2-20200127
- ▶ Also on GitHub ThaleMine and ThaleMine Bio Sources

# BAR ThaleMine: What's working

- ▶ data categories figure



# BAR ThaleMine: Some future plans and Known Issues

The follow is planned

- ▶ Arabidopsis Strains
- ▶ Gene expression data (Araport11 RNA-Seq)
- ▶ Salk data (TDNA-Seq)
- ▶ Gene histories
- ▶ TAIR annotations

Known Issues

- ▶ web properties from ThaleMine web service
- ▶ Bluegene frontend

# Acknowledgements