ENABLING INTEGRATIVE MODELING OF HUMAN IMMUNOLOGICAL DATA WITH IMMUNESPACE

RAPHAEL GOTTARDO, PHD
VACCINE AND INFECTIOUS DISEASE DIVISION
FRED HUTCH

ImmuneSpace

suppported by
Human Immunology Project Consortium

THE HIPC CONSORTIUM
HISTORY AND OVERVIEW

- The human immunology project consortium

  - Established in 2010 by the NIAID Division of Allergy, Immunology, and Transplantation

  - Collaborative effort across 7 centers

- Characterize the status of the immune system in diverse populations under both normal conditions and in response to stimuli

  - Generate large amount of data (including high-dimensional data)

  - Use systems biology approaches

ImmuneSpace
suppported by
HIPC Human Immunology Project Consortium

IMMUNESPACE
FROM DATA TO KNOWLEDGE

- High-quality public web-interface to HIPC data, analysis results and tools

- Build using LabKey server, inspired from TrialShare but customized to fit our needs

- Data and analysis engine complementary to ImmPort

- Integrate HIPC funded bioinformatics tools

- Enable cross-assay, cross-study integration

- Public resource for immunology (public face of the consortium)

ImmuneSpace

HIPC supppported by
Human Immunology Project Consortium
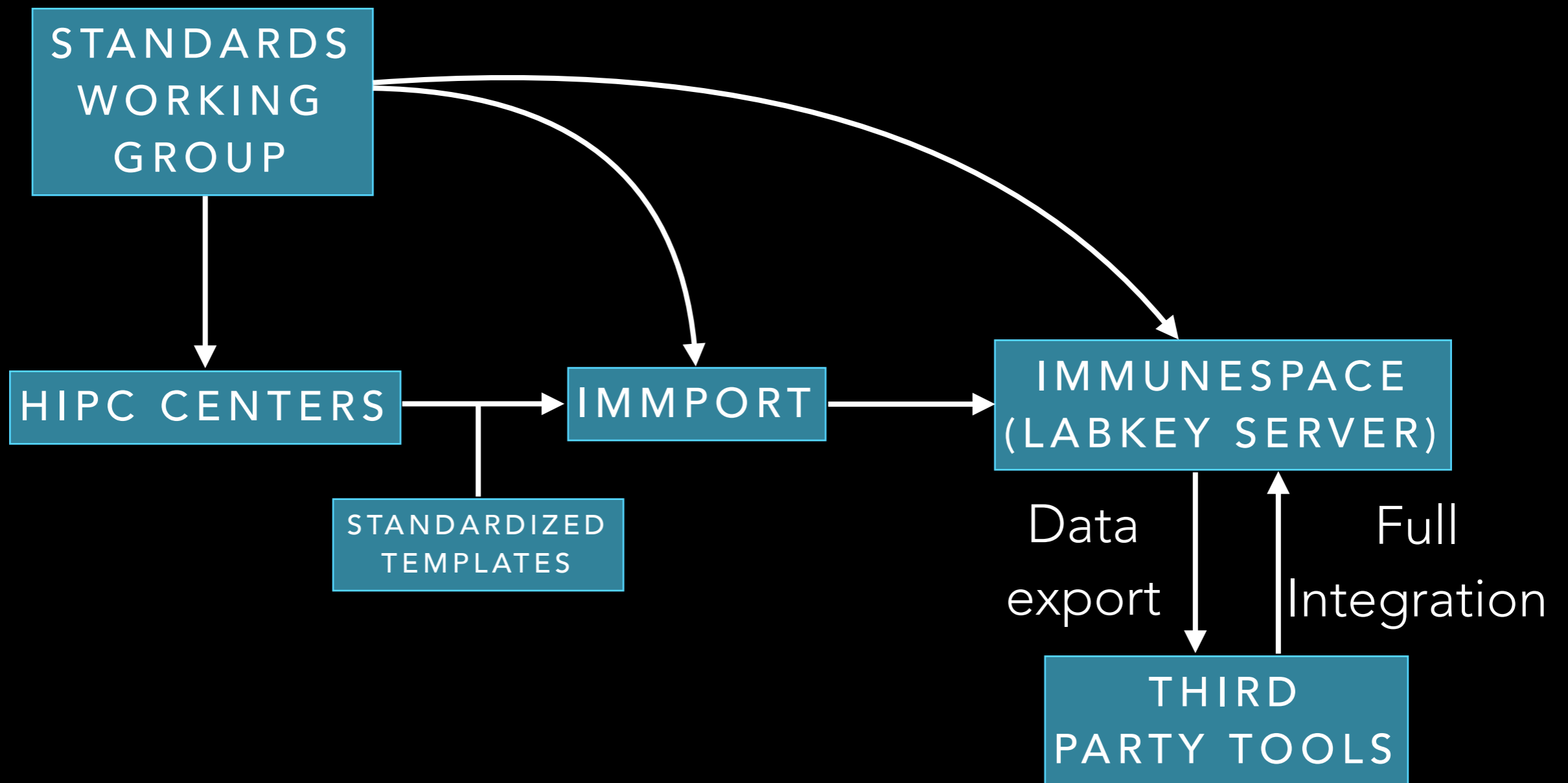
R AND BIOCONDUCTOR
GOLD STANDARD FOR DATA ANALYSIS

- Open-source software for statistical computing and graphics

- State-of-the-art for data analysis and data visualization

- De-facto standard for data analysis in Bioinformatics thanks to Bioconductor

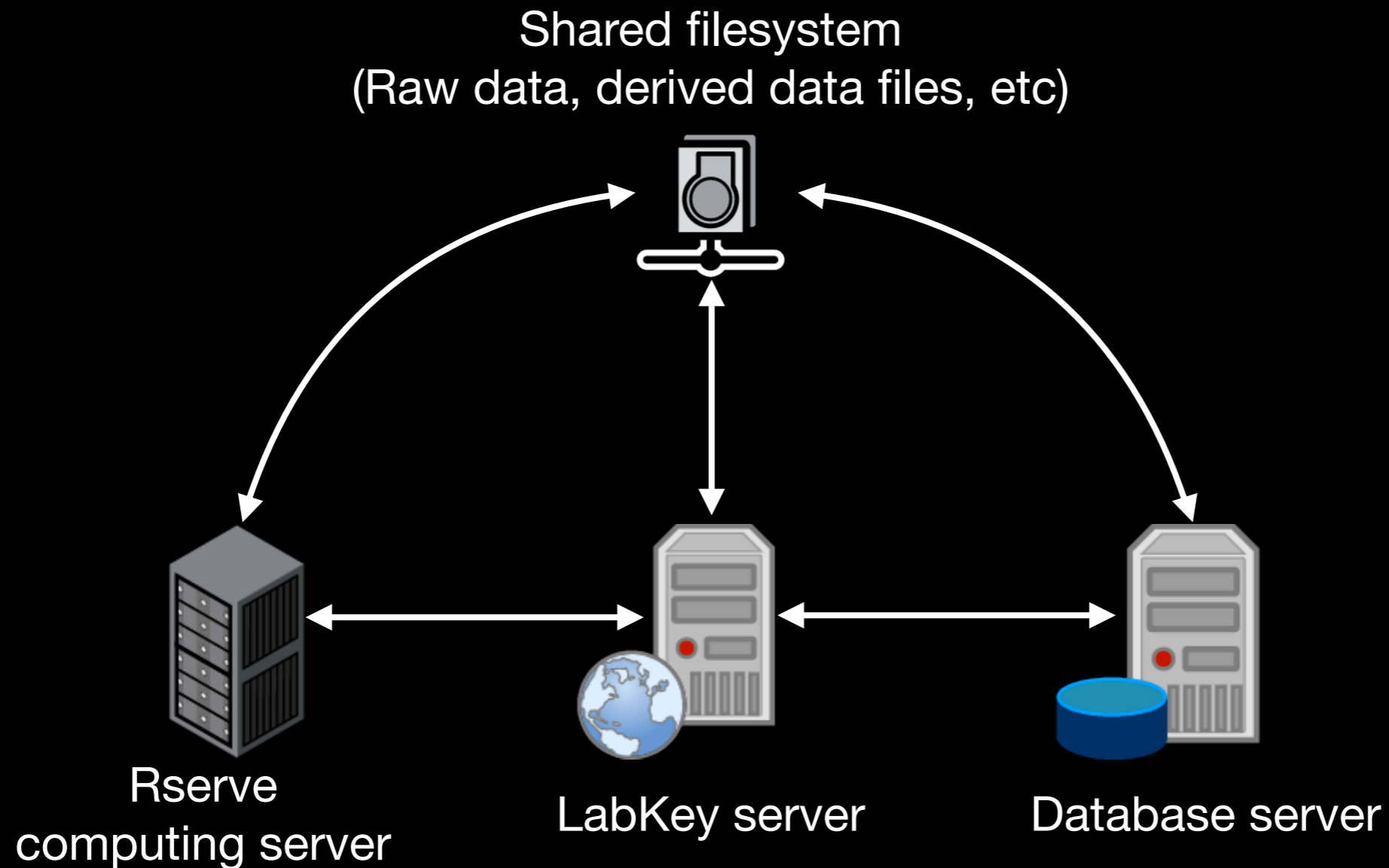- Community driven, with over 5000 packages available to extend the base functionalities

ImmuneSpace
suppported by
HIPC Human Immunology Project Consortium

5

LEVERAGING R AND BIOCONDUCTOR
NEW R↔LABKEY INTERFACES

- **Rserve (TCP/IP interface to R):** consistent R session, session management, remote server, etc.

- **knitr:** dynamic report generation

  - Enhanced R report

  - Mix and match computer code with regular text and figures

  - Efficient caching mechanism for large computation

  - Fully reproducible!

ImmuneSpace

suppported by
HIPC Human Immunology Project Consortium

IMMUNESPACE
INFRASTRUCTURE AND SETUP

Shared filesystem
(Raw data, derived data files, etc)

Rserve
computing server

LabKey server

Database server

ImmuneSpace

COMBINING R AND LABKEY
THE BEST OF BOTH WORLDS

- **Standardized analysis modules in ImmuneSpace**

  - Gene expression data pre-processing

  - Data visualization leveraging *ggplot2*

  - Differential gene expression and gene set enrichment analyses

  - Multivariate prediction

ImmuneSpace

supppported by
Human Immunology Project Consortium

DEMO 1: THE STUDY FINDER

USING THE STUDY FINDER TO FIND INTERESTING DATA

ImmuneSpace
supppported by
HIPC Human Immunology Project Consortium

DEMO 2: THE DATA EXPLORER MODULE

STANDARDIZED DATA EXPLORATION AND VISUALIZATION

ImmuneSpace
suppported by
Human Immunology Project Consortium

PREDICTING INFLUENZA VACCINATION RESPONSES FROM GENE EXPRESSION

(Time Permitting)

ImmuneSpace
supported by
HIPC Human Immunology Project Consortium

IMMUNESPACER
SIMPLIFYING ACCESS TO DATA FROM R

- A thin wrapper around the Rlabkey package

- Simplified query syntax

- Download selected datasets to the R session as annotated BioConductor objects (e.g. ExpressionSets)

- Based on the *S5* reference classes for increased efficiency (e.g. memory, caching, etc)

ImmuneSpace
supported by
HIPC Human Immunology Project Consortium

IMMUNESPACER VS. RLABKEY
COMPARING SYNTAX

```
> study <- CreateConnection("SDY269")
> study$listDatasets()
      cohort_membership
      demographics
      hai
      fcs_analyzed_result
      elispot
      elisa
      pcr
      fcs_files
      gene_expression_files
Expression Matrices
      LAIV_2008
      TIV_2008
> study$getDataset("hai")
```

ImmuneSpaceR

```
> labkey.selectRows(baseUrl = "http://www.immunespace.org",
folderPath = "/Studies/SDY269", schemaName = "study",
queryName = "hai")
```

Rlabkey

ImmuneSpace

supported by
HIPC Human Immunology Project Consortium

AND SOME TOPICS FOR DISCUSSION…

- Enable cross-study data integration within ImmuneSpace (generalize our current modules)

- Add novel analysis tools and study data

- Support novel assay data and technologies

- Community driven repository for modules?

  - Standardized API and guidelines

ImmuneSpace
supported by
HIPC Human Immunology Project Consortium

15

ACKNOWLEDGMENTS
A TEAM EFFORT

- **HIPC members:** Vladimir Brusic, Steve Kleinstein, Ann Oberg, Mark Davis, Peter Li, Damien Chaussabel, Nick Haining, Shai Shen-Orr, Ashley Xia

- **ImmPort team:** Jeff Wiser, and Patrick Dunn

- **LabKey:** Matt Bellew, Kevin Krouse, Kristin Fitzsimmons, Brian Connolly, Dax Hawkins

- **FHCRC:** Renan Sauteraud, Greg Finak, Lev Dashevkiy, and Erik Swanson

- **ITN:** Adam Asare, Tanya Antonille

- **Gates Foundation:** CDS project

- **HIPC steering committee**

- **NIH/NIAID:** Helen Quill, Alkis Togias, Ashley Xia, Dawei Lin, Quan Chen

ImmuneSpace
supppported by
HIPC Human Immunology Project Consortium

Funded by:

National Institute
of Allergy and
Infectious Diseases