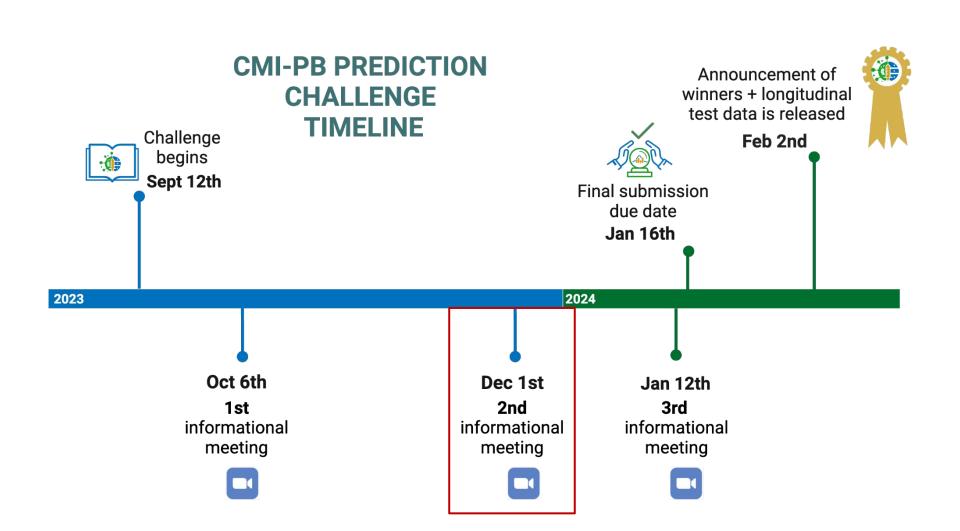
CMI-PB Prediction Challenge

2nd Informational Session December 1st, 2023

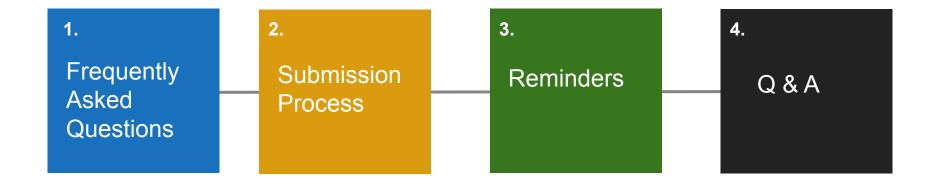
La Jolla Institute for Immunology





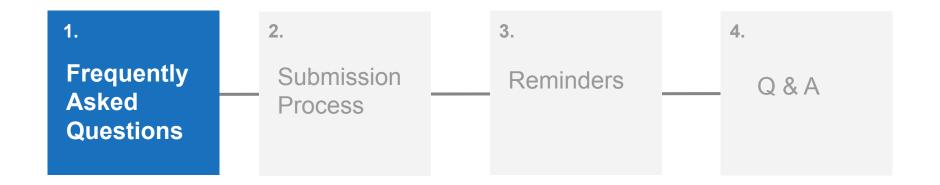


Agenda for Today's Session





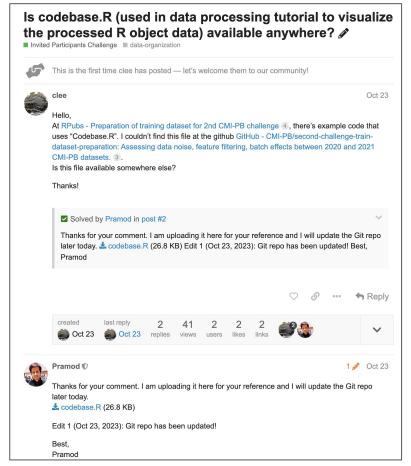






- We added all the codebase related data processing pipeline as GitHub <u>repo</u> as well.
- To find a more comprehensive discussion, access the Solutions Center here.

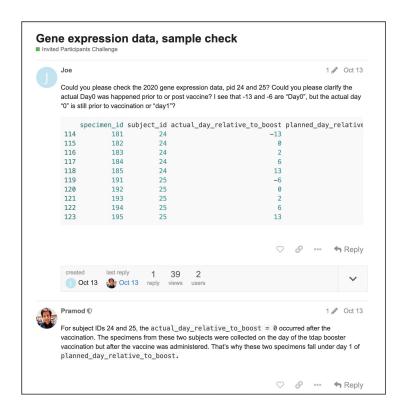








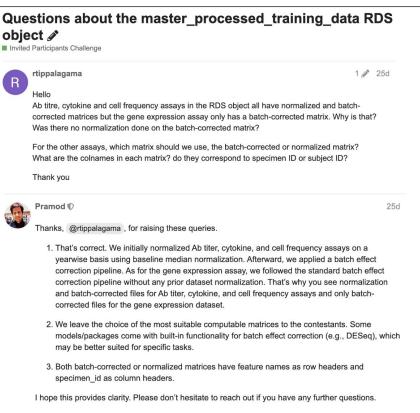
- When actual_day_relative_to_boost is 0, the sample was collected from the donor on the same day, but after, the tdap booster was administered.
- To find a more comprehensive discussion, access the Solutions Center <u>here</u>.



3. Master processed training data



- Ab titer, cytokine, and cell frequency assays are normalized on a yearwise basis and a batch effect correction pipeline is applied afterward.
- For the gene expression assay, only the standard batch effect correction pipeline is applied.
- To find a more comprehensive discussion, access the Solutions Center <u>here</u>.



4. CMI-PB Challenge Data Set tracking



- Participants have noticed inconsistencies and minor issues with the dataset which has prompted us to change some information.
- The current dataset can be found <u>here</u> or on the CMI-PB site in the Prediction Challenge section under Data and Resources.
- To follow updates and find a more detailed conversation about the legacy of datasets, please refer to this discussion on the Solutions Center.

2nd Challenge Dataset changes tracking

■ Invited Participants Challenge ■ website ■ api-access



Pramod ®

4 @ Oct 13

As the 2nd challenge progresses, contestants might notice inconsistencies or issues in the dataset. This might result in challenge datasets to undergo modifications over time. This page is dedicated to organizing and tracking all changes related to the datasets. Older (legacy) versions of the dataset are stored in the legacy repository, and the updated (current) datasets are available here 8.

2023 Sept 12

- Datasets are made available to 2nd challenge contestants via API and direct download 8.
- . This version of the dataset can be found in the legacy repository here.

2023 Oct 10 (current dataset version)

- . Few contestants reported issues when accessing the data files. The identified issues include
- Inconsistencies in the actual dates relative to the boost.** A more detailed discussion on this can be found here
- The names of cell populations in the prediction dataset differed from those in the training dataset **
- A more detailed discussion on this can be found here 1
- This is the current version of the challenge dataset and is accessible here 8.



Joe

Oct 13

could you please check file "2020LD_pbmc_cell_frequency.tsv" in this path (cmipb_challenge_datasets/current/2nd_challenge/raw_datasets/training_data/), file size indicates not all cell fred data included in the file. See this below:

unique(cellfreg 2020\$cell type name)

- [1] "Monocytes" "CD33HLADR" "Classical_Monocytes" "Non-Classical_Monocytes" "Intermediate Monocytes" "Bcells"
- [7] "CD3CD19" "CD3CD19neg" "CD3 Tcells" "CD4Tcells" "CD8Tcells" "Tregs"
- [13] "TemraCD4" "NaiveCD4" "TemCD4" "TcmCD4" "TemraCD8" "NaiveCD8"
 - [19] "TemCD8" "TcmCD8" "NK" "Basophils" "mDC" "pDC"
- [25] "ASCs (Plasmablasts)"



Pramod ®

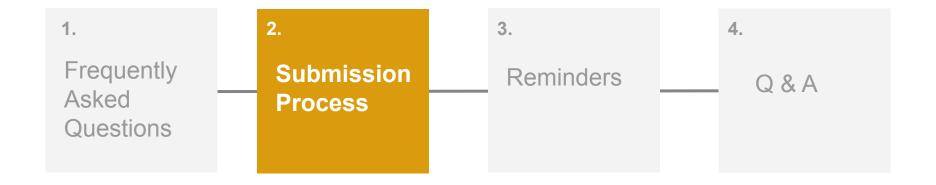
Oct 13

Hi @Joe 2020 cohort has different feature counts when compared to 2020 and 2021 datasets. The 2020 cohort's cell frequency dataset comprises 25 cell types, as you pointed out., while the 2021 and 2022 datasets each have 50 cell types.

Similarly, the 2020 plasma_cytokine_concentration (Olink) data has 263 cytokines, while the 2021 and 2022 plasma_cytokine_concentration datasets each have 45 cytokines. Genes in all three datasets are identical. I hope this information is helpful.



Agenda for Today's Session





Final submissions December 1, 2023 - January 16, 2024

 Note: Submissions during the preliminary submissions phase (September 12 - November 30) will not be considered a final submission. You must <u>re-submit</u> your entry at any point from December 1, 2023 to January 16, 2024 to be entered into the challenge.



Submission Process Demonstration

https://www.cmi-pb.org/

- 1) Manual data entry using Google sheets
- 2) Using coding pipeline

rcl

ign Out

×

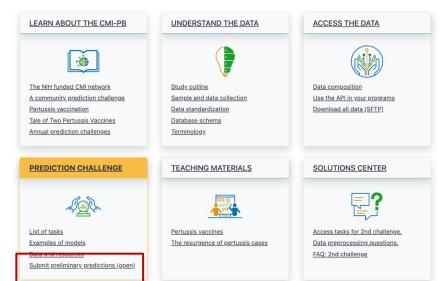


2nd CMI-PB Prediction Challenge

Revolutionizing computational modelling approach for immune response prediction

Learn more: Training data, Prediction tasks, 1st challenge | Participate

The mission of CMI-PB is to provide the scientific community with a comprehensive, high-quality and freely accessible resource of Pertussis booster vaccination.





6	378		
Antibody levels	~		Search

Home > 2nd CMI-PB Prediction challenge: Overview > Prediction challenge Submission

2nd CMI-PB Prediction challenge Submission

Table of contents

- · Prepare submission file
- · Make preliminary submission
- · Access past submissions

Welcome to the 2nd CMI-PB Prediction challenge. Please follow the steps below to submit your prediction challenge. Currently, we have If you have any issues, use our solutions center to post your questions.

Step 1: Prepare Submission File

Create a model and complete your analysis. We only accept submissions in the given Tab-separated values (TSV) file format:

- 1. Download the submission template
- 2. Submit your prediction in the prescribed format.

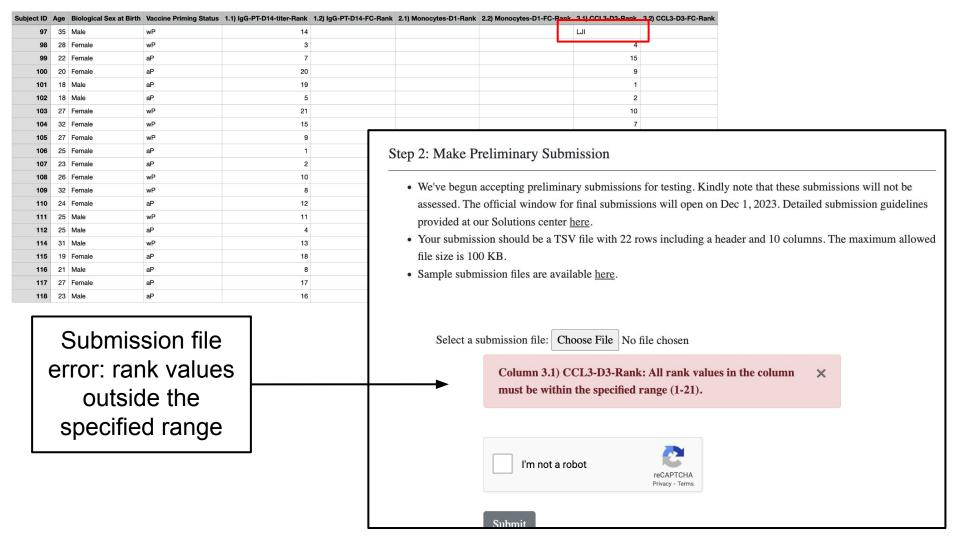
Step 2: Make Preliminary Submission

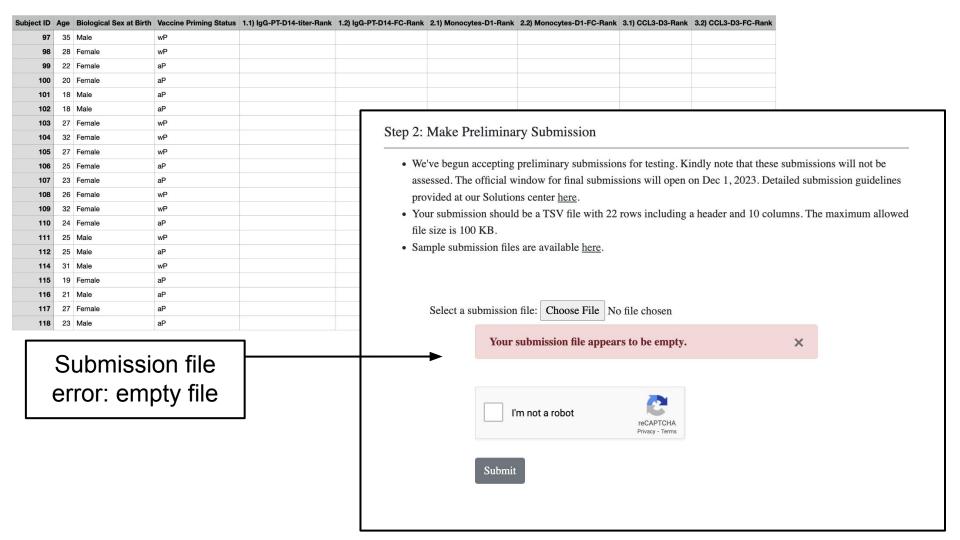
- We've begun accepting preliminary submissions for testing. Kindly note that these submissions will not be
 assessed. The official window for final submissions will open on Dec 1, 2023.
- Your submission should be a TSV file with 22 rows including a header and 10 columns.
- · Sample submission files are available here.

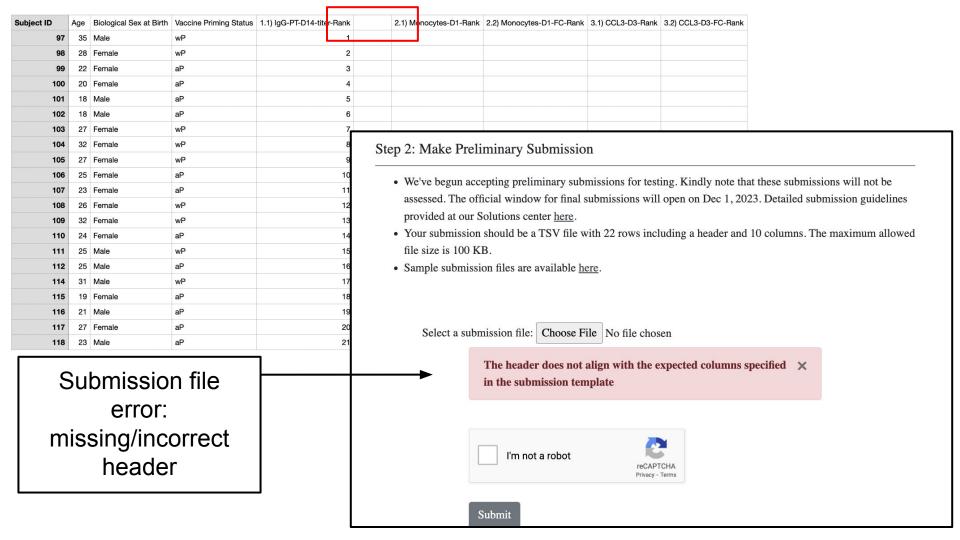


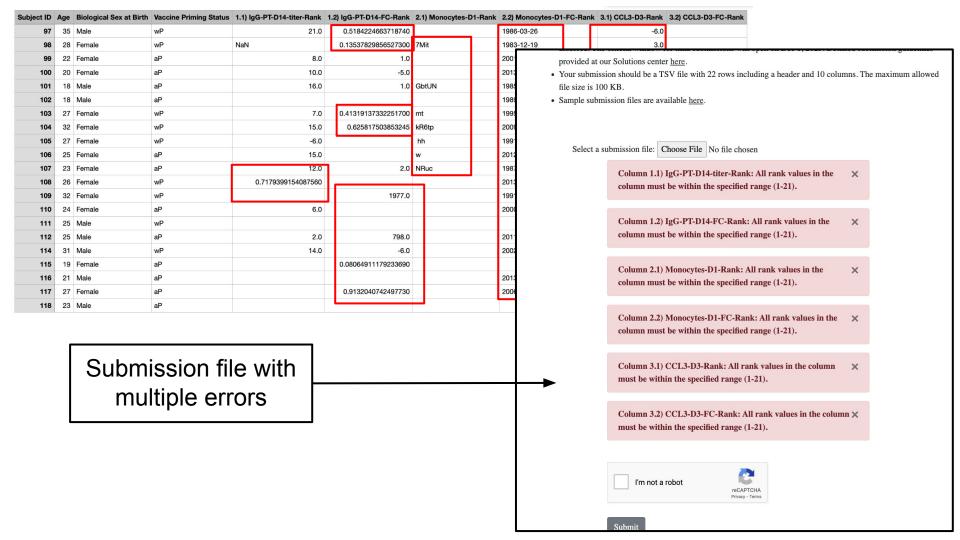


Submission File Errors











Successful Submissions



Submission with all tasks

Subject ID	Age	Biological Sex at Birth	Vaccine Priming Status	1.1) IgG-PT-D14-titer-Rank	1.2) IgG-PT-D14-FC-Rank	2.1) Monocytes-D1-Rank	2.2) Monocytes-D1-FC-Rank	3.1) CCL3-D3-Rank	3.2) CCL3-D3-FC-Rank
97	35	Male	wP	14	6	15	21	11	17
98	28	Female	wP	3	7	13	15	4	1
99	22	Female	аР	7	2	18	12	15	13
100	20	Female	аР	20	20	10	11	9	6
101	18	Male	аР	19	5	1	8	1	19
102	18	Male	аР	5	21	11	17	2	9
103	27	Female	wP	21	4	4	7	10	12
104	32	Female	wP	15	12	14	16	7	14
105	27	Female	wP	9	15	17	13	20	11
106	25	Female	аР	1	11	3	18	19	3
107	23	Female	аР	2	16	19	10	5	4
108	26	Female	wP	10	17	7	1	21	5
109	32	Female	wP	8	18	12	20	8	15
110	24	Female	аР	12	13	16	19	12	2
111	25	Male	wP	11	8	20	5	14	16
112	25	Male	аР	4	10	5	9	3	20
114	31	Male	wP	13	1	21	14	13	8
115	19	Female	аР	18	3	8	2	18	21
116	21	Male	аР	8	19	8	4	8	7
117	27	Female	аР	17	9	9	8	17	18
118	23	Male	аР	16	14	2	3	16	10

*Populated with random numbers

Submission with two tasks



Subject ID	Age	Biological Sex at Birth	Vaccine Priming Status	1.1) IgG-PT-D14-titer-Rank	1.2) IgG-PT-D14-FC-Rank	2.1) Monocytes-D1-Rank	2.2) Monocytes-D1-FC-Rank	3.1) CCL3-D3-Rank	3.2) CCL3-D3-FC-Rank
97	35	Male	wP	14				11	
98	28	Female	wP	3				4	
99	22	Female	aP	7				15	
100	20	Female	аР	20				9	
101	18	Male	aP	19				1	
102	18	Male	aP	5				2	
103	27	Female	wP	21				10	
104	32	Female	wP	15				7	
105	27	Female	wP	9				20	
106	25	Female	aP	1				19	
107	23	Female	aP	2				5	
108	26	Female	wP	10				21	
109	32	Female	wP	8				8	
110	24	Female	aP	12				12	
111	25	Male	wP	11				14	
112	25	Male	aP	4				3	
114	31	Male	wP	13				13	
115	19	Female	aP	18				18	
116	21	Male	аР	8				8	
117	27	Female	aP	17				17	
118	23	Male	aP	16				16	

*Populated with random numbers



Antibody levels ~		Search
-------------------	--	--------

2nd CMI-PB Prediction challenge Submission

Your submission has been recorded successfully. A confirmation email has been sent to you. Thank you!

Useful links:

- Past submissions
- · CMI-PB home

Please let us know if you have any questions at our CMI-PB solutions center.

GET HELP

About Us

Solutions Center

1000

PUBLICATIONS

Latest build: September 2023 Version history

Version history
APIs

DATASETS

Downloads

BioRxiv 2023 JCI Insights 2021 La Jolla Institute FOR IMMUNOLOGY Life Without Disease.

Supported by a grant from the National Institutes of Health (NIH U01 AI150753) | ©2021-2023: La Jolla Institute for Immunology





WELCOME TO THE 2ND CMI-PB PREDICTION CHALLENGE

Dear sorfield@lji.org,

submission@cmi-pb.org via lji.org

to me, aazhan 🕶

Thank you for submitting your response to the 2nd CMI-PB Prediction challenge. Your response is now successfully recorded.

The submitted file is attached here for your reference. Feel free to access all your past submissions here. If you ever change your mind and want to re-submit, please make sure to enter all your answers in the sheet again as your new submission overrides all previous submissions.

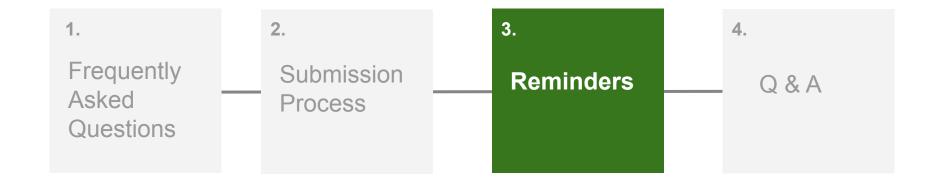
We look forward to reviewing your results! In the meantime, please let us know if you have any questions at our CMI-PB solutions center.

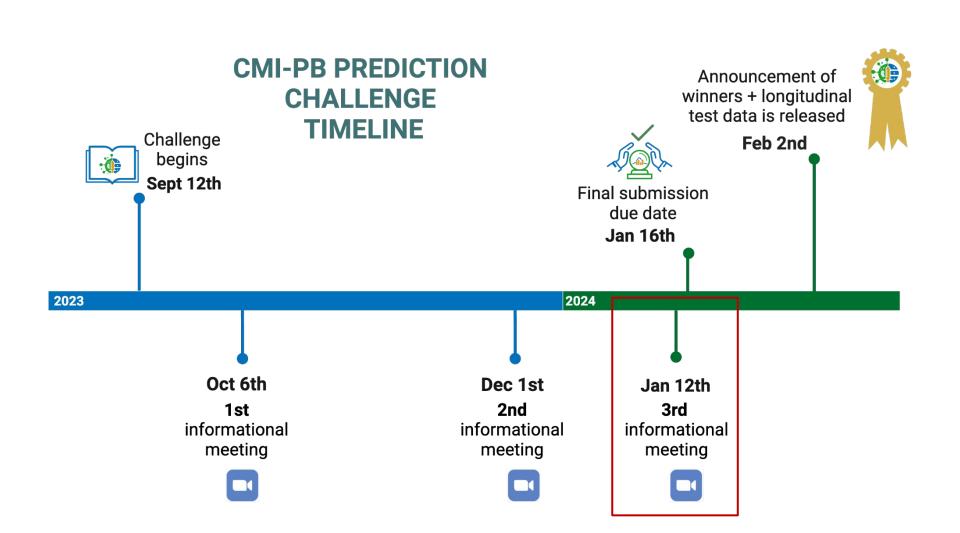
Best wishes. CMI-PB Team

You will receive an email confirmation once your submission is received.



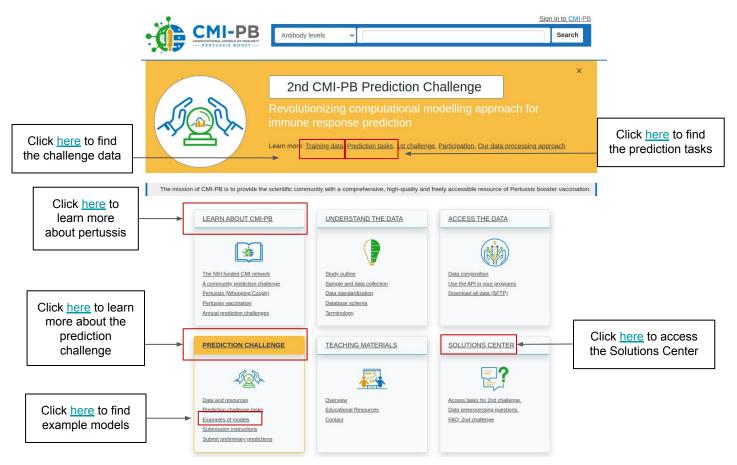






Other resources on the site





The CMI-PB team



Kleinstein Lab (Yale)



- Expertise: A combination of "big data" analysis and immunology domain.
- Collaborating on data and models being released to the community to support reproducibility and the prediction contest, and also participate in the prediction challenge.

Steven Kleinstein Jeremy Gygi Leying Guan Anna Konstorum

Grant Lab (UCSD)



- Expertise: the use of computational approaches, based on both biophysics and bioinformatics, to study the structure, function and evolution of key biological macromolecules.
- Dr. Grant will engage and advise over 40 biology graduate students in the CMI-PB Prediction Challenge.

Barry Grant

Ay Lab (LJI)



- Expertise: Development of bioinformatics tools that utilize high-dimensional and high-throughput datasets to deduce insights into chromatin conformation, genetic variation, and the regulation of gene expression.
- The Ay lab is focused on developing predictive machine learning models, which will serve as examples and baselines for participants in the CMI-PB challenge.

Ferhat Ay Joaquin Reyna

Peters Lab (LJI)



- Expertise: Both experimental and computational studies to better understand human immune responses in the context of infectious diseases, allergy, cancer and vaccines.
- The team is responsible for the generation of experimental data, making it accessible in a central and standardized fashion, and coordinating the creation and coordination of the prediction contest.

Bjoern Peters Jason Greenbaum James Overton Brendan Ha Pramod Shinde Mari Kojima Rasteh Haji Kazem Nili Jiyeun Lee Lisa Willemsen Shelby Orfield

And thank you to the Sette Lab, Crotty lab, LJI Clinical Core, LJI Bioinformatics Core

The CMI-PB team members

















Steven Kleinstein

Ferhat Ay

Barry Grant

Shane Crotty

Alessandro Sette







Shelby Orfield



Lisa Willemsen



Leying Guan



Joaquin Reyna



Mari Kojima



Ferran Soldevila



Rasteh Nili



Jason Greenbaum



Brendan Ha



Jiyeun Lee



Ricardo De Silva Antunes



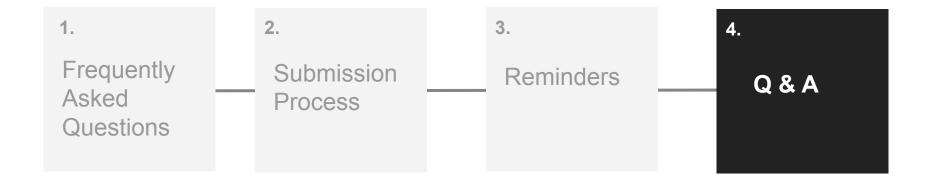
Jeremy Gygi



Anna Konstorum



Agenda for Today's Session





Questions?

Please email cmi-pb-challenge@lji.org at any time throughout the challenge with any questions.