

Work in Progress: *Bioconductor* and the HCA

Martin.Morgan@RoswellPark.org

20 February 2019

Scope

Funding

- One year seed funding
- 8 Projects
- Fully 'open' development, initiated by Aaron Lun

Aims

1. HCA Access
2. Data representations
3. Scalable preprocessing pipelines
4. Scalable statistical routines
5. Ontology

Symposium

- Near end of seed funding, so projects are approaching maturity
- Great opportunity to
 - Summarize progress
 - Identify challenges & opportunities
 - Share with a broader community!

'Work in progress'

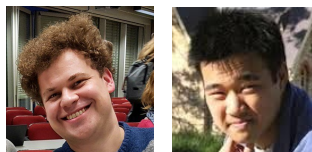
- Rough and challenging edges, particularly in scalability

People and collaborations

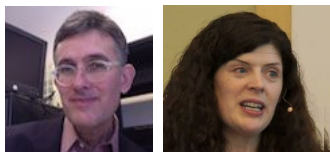
Data access and annotation



Martin Morgan, Daniel van Twisk, Marcel Ramos



John Marioni, Aaron Lun



Vince Carey, Aedin Culhane



Raphael Gottardo, Mike Jiang, Greg Finak

Scalable interactivity in R



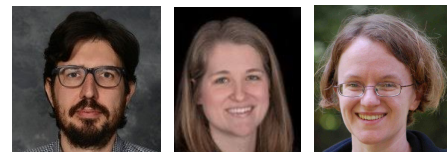
Wolfgang Huber, Mike Smith



Scalable statistical procedures



Kasper Hansen, Peter Hickey



Davide Risso, Stephanie Hicks, Elizabeth Purdom



Rafael Irizarry, Christina Kendziorski, Keegan Korthauer, Meromit Singer

Acknowledgements (Martin)



National Human Genome
Research Institute

NATIONAL CANCER INSTITUTE
Informatics Technology for
Cancer Research



A portion of this work is supported by the Chan Zuckerberg Initiative DAF, an advised fund of Silicon Valley Community Foundation.

Research reported in this presentation was supported by the NHGRI and NCI of the National Institutes of Health under award numbers U41HG004059, U24CA180996, and U24CA232979. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

This work was performed on behalf of the SOUND Consortium and funded under the EU H2020 Personalizing Health and Care Program, Action contract number 633974.