

Disclaimer: This is a machine generated PDF of selected content from our databases. This functionality is provided solely for your convenience and is in no way intended to replace original scanned PDF. Neither Cengage Learning nor its licensors make any representations or warranties with respect to the machine generated PDF. The PDF is automatically generated "AS IS" and "AS AVAILABLE" and are not retained in our systems. CENGAGE LEARNING AND ITS LICENSORS SPECIFICALLY DISCLAIM ANY AND ALL EXPRESS OR IMPLIED WARRANTIES, INCLUDING WITHOUT LIMITATION, ANY WARRANTIES FOR AVAILABILITY, ACCURACY, TIMELINESS, COMPLETENESS, NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Your use of the machine generated PDF is subject to all use restrictions contained in The Cengage Learning Subscription and License Agreement and/or the Gale Virtual Reference Library Terms and Conditions and by using the machine generated PDF functionality you agree to forgo any and all claims against Cengage Learning or its licensors for your use of the machine generated PDF functionality and any output derived therefrom.

A Comprehensive Infrastructure for Big Data in Cancer Research: Accelerating Cancer Research and Precision Medicine

Izumi V. Hinkson, Tanja M. Davidsen, Juli D. Klemm, Anthony R. Kerlavage and Warren A. Kibbe

(Sept. 21, 2017):

DOI: <http://dx.doi.org/10.3389/fcell.2017.00083>

Copyright: COPYRIGHT 2017 Frontiers Research Foundation

<http://journal.frontiersin.org/journal/cell-and-developmental-biology>

Author(s): Izumi V. Hinkson, Tanja M. Davidsen, Juli D. Klemm, Anthony R. Kerlavage, Warren A. Kibbe

Advancements in next-generation sequencing and other -omics technologies are accelerating the detailed molecular characterization of individual patient tumors, and driving the evolution of precision medicine. Cancer is no longer considered a single disease, but rather, a diverse array of diseases wherein each patient has a unique collection of germline variants and somatic mutations. Molecular profiling of patient-derived samples has led to a data explosion that could help us understand the contributions of environment and germline to risk, therapeutic response, and outcome. To maximize the value of these data, an interdisciplinary approach is paramount. The National Cancer Institute (NCI) has initiated multiple projects to characterize tumor samples using multi-omic approaches. These projects harness the expertise of clinicians, biologists, computer scientists, and software engineers to investigate cancer biology and therapeutic response in multidisciplinary teams. Petabytes of cancer genomic, transcriptomic, epigenomic, proteomic, and imaging data have been generated by these projects. To address the data analysis challenges associated with these large datasets, the NCI has sponsored the development of the Genomic Data Commons (GDC) and three Cloud Resources. The GDC ensures data and metadata quality, ingests and harmonizes genomic data, and securely redistributes the data. During its pilot phase, the Cloud Resources tested multiple cloud-based approaches for enhancing data access, collaboration, computational scalability, resource democratization, and reproducibility. These NCI-led efforts are continuously being refined to better support open data practices and precision oncology, and to serve as building blocks of a future NCI Cancer Research Data Commons.

To access, purchase, authenticate, or subscribe to the full-text of this article, please visit this link:

<https://www.frontiersin.org/articles/10.3389/fcell.2017.00083>

Source Citation (MLA 8th Edition)

Hinkson, Izumi V., et al. "A Comprehensive Infrastructure for Big Data in Cancer Research: Accelerating Cancer Research and Precision Medicine." *Frontiers in Cell and Developmental Biology*, 2017. *Academic OneFile*, go.galegroup.com/ps/i.do?p=AONE&sw=w&u=duke_perkins&v=2.1&id=GALE%7CA511188704&it=r&asid=187f716548b2a678dc7741d79a427e4e. Accessed 30 Oct. 2017.

Gale Document Number: GALE|A511188704
