

# NCBI Pathogen Detection Update

Arjun Prasad



U.S. National Library of Medicine  
National Center for Biotechnology Information

If You Remember One Thing...

**pd-help@ncbi.nlm.nih.gov**

And if you can remember another thing

**www.ncbi.nlm.nih.gov/pathogens**

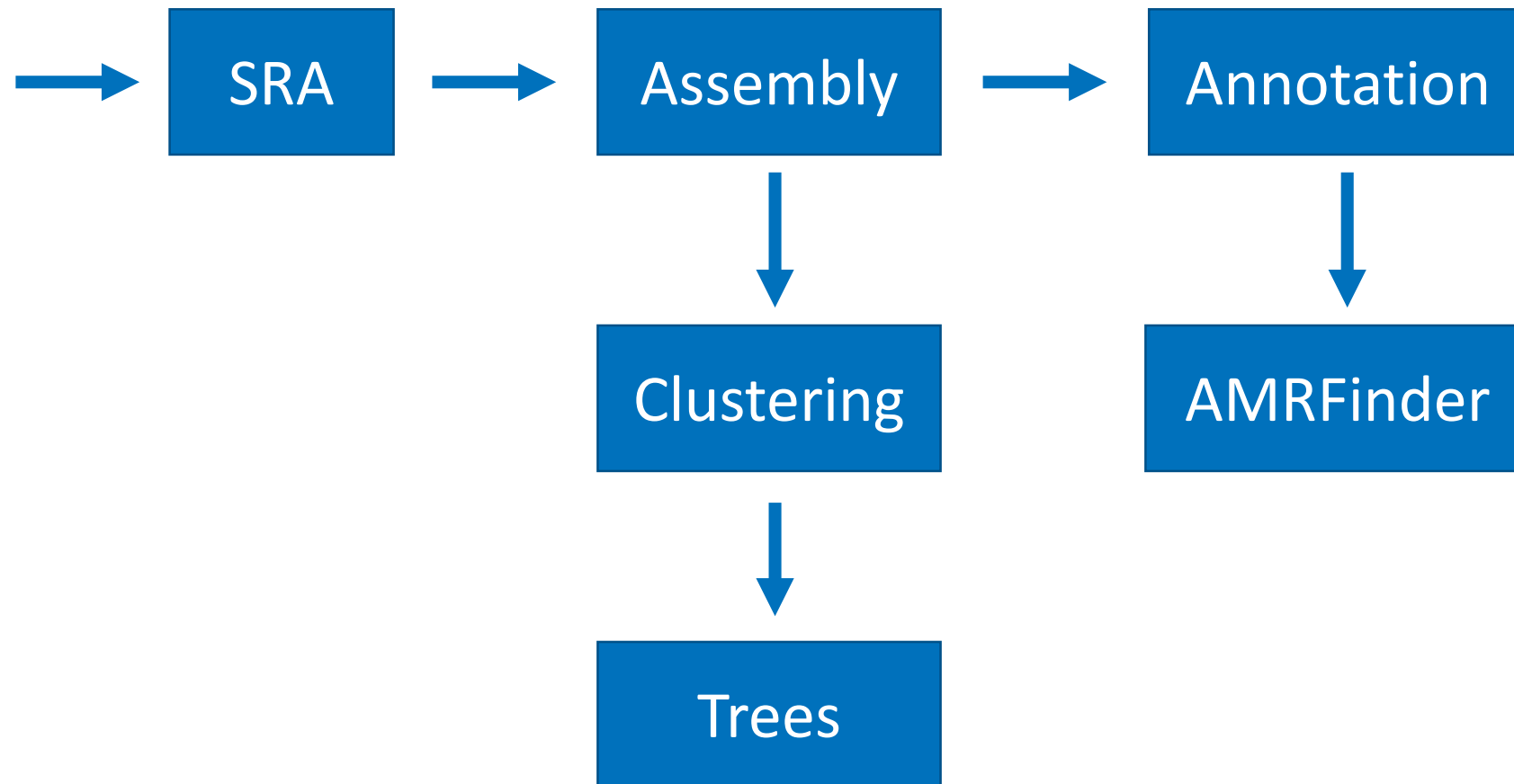
# Pathogen Detection Analysis goals

1. Are these isolates clonally related?
2. What is the anti-microbial resistance gene repertoire of this isolate?

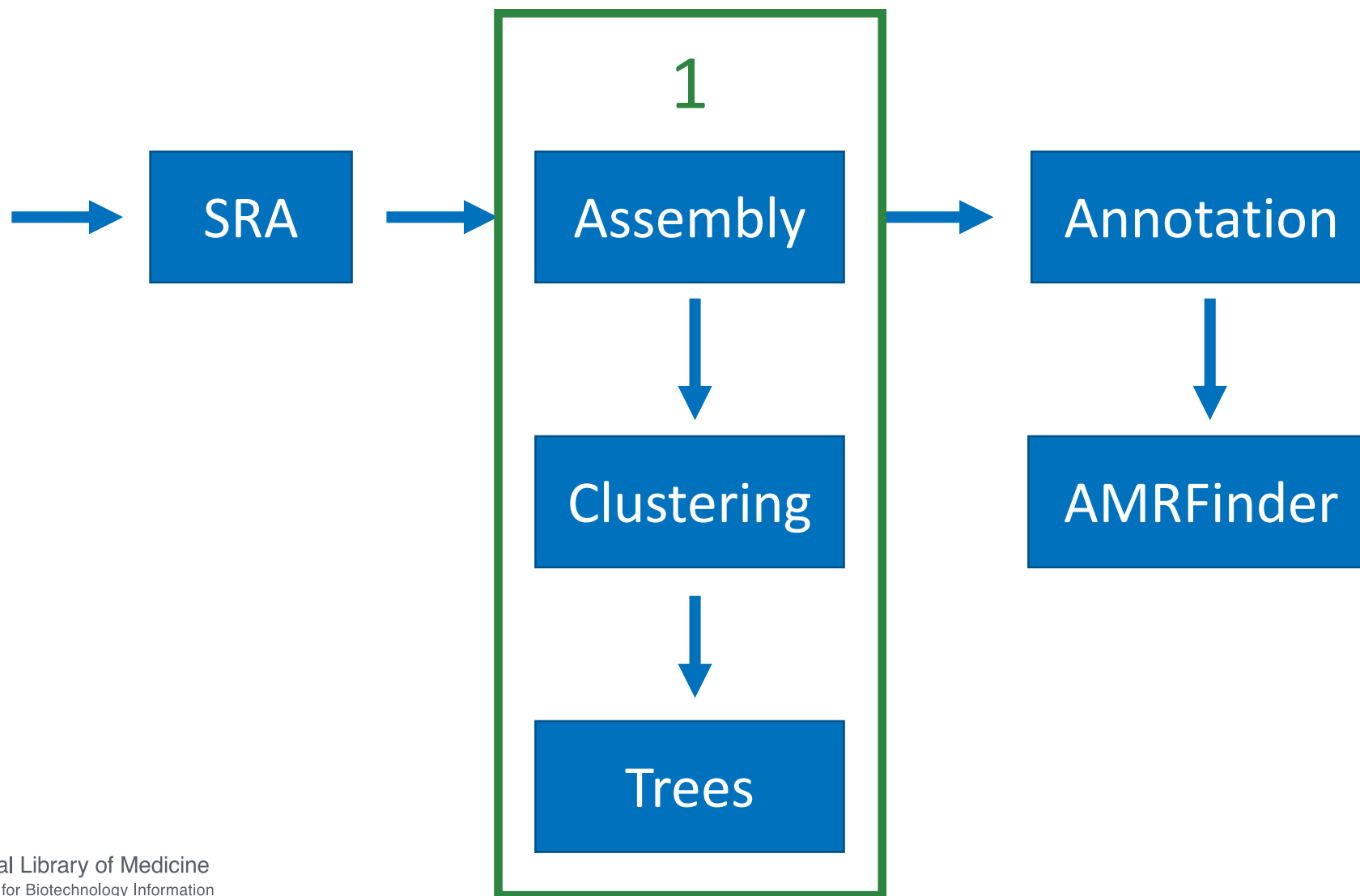
# How do we do this?

- Collaborate with public health agencies and encourage people to submit
  - Isolate WGS or assemblies
  - Currently genome sequences for over 275,000 isolates
  - Expect ~90,000 new US isolates next year
- Assemble genomes and analyze
  - Analyze the sequences by comparing to others
  - Annotate the assemblies and look for AMR genes

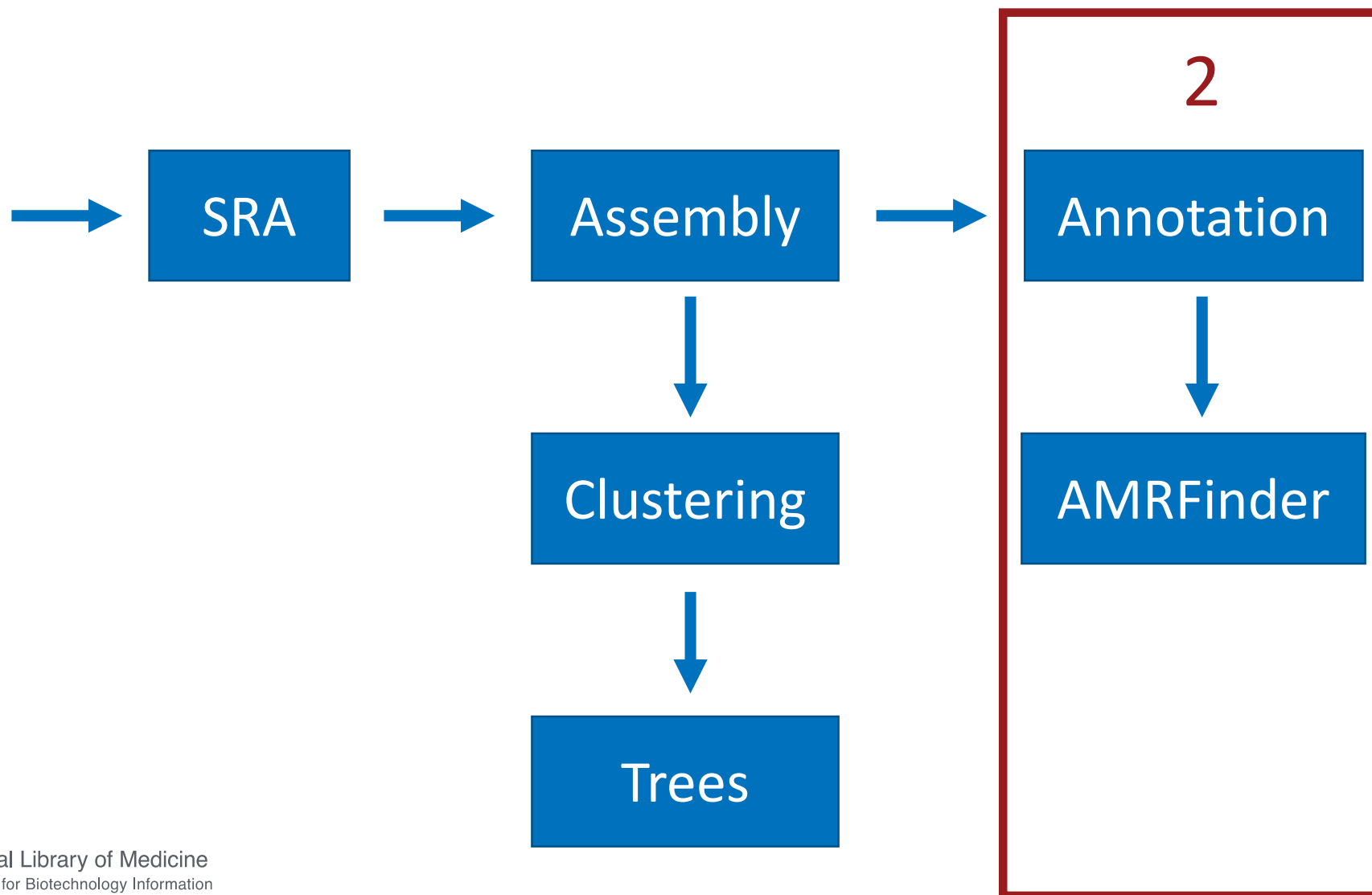
# Pathogen Detection Pipeline



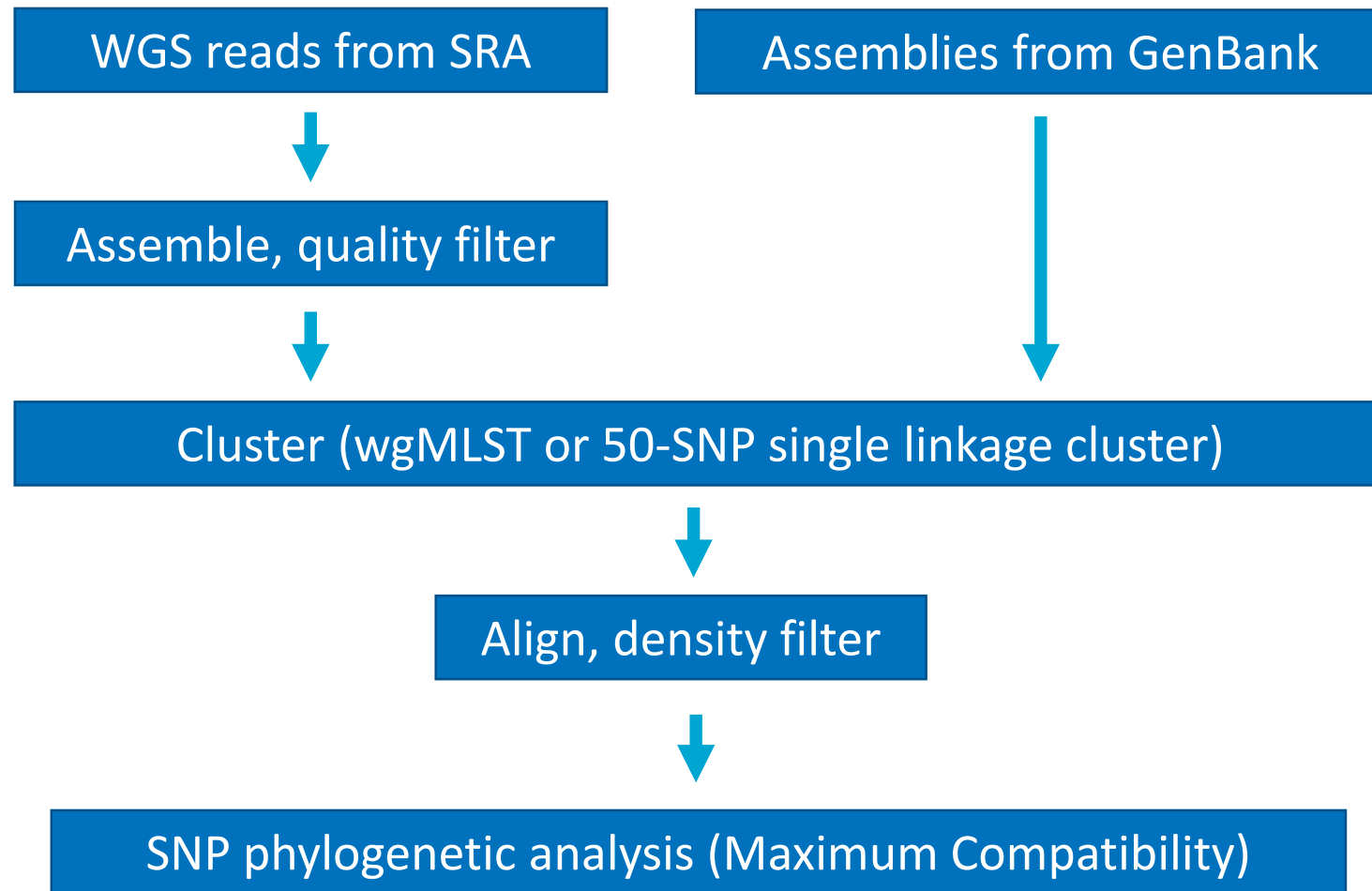
# Are these isolates clonally related?



# What are the AMR genes?



# 1. Are these isolates clonally related?





## 2. What is the anti-microbial resistance gene repertoire of this isolate?

- Curated and released a hierarchical database of acquired anti-microbial resistance (AMR) genes and HMMs
- Created software (AMRFinder) to identify AMR proteins
- Created database to accept and store antibiograms to associate with sequences

# On to the demo...

- <http://ncbi.nlm.nih.gov/pathogens>

- New: <https://www.ncbi.nlm.nih.gov/pathogens/isolates/#/search/new:1>
- Flour: <https://www.ncbi.nlm.nih.gov/pathogens/isolates#/search/flour>
  - treelabel=epi\_type,isolation\_source,biosample\_acc:  
[https://www.ncbi.nlm.nih.gov/Structure/tree/#!/cluster/Escherichia\\_coli\\_Shigella/PDG00000004.1015/PDS000003441.82/70?key=C0CkKbge1dqijGXtyu15iOMmL04wcuZpItDvoYXB15ue2NaXlfba?treelabel=epi\\_type,isolation\\_source,biosample\\_acc](https://www.ncbi.nlm.nih.gov/Structure/tree/#!/cluster/Escherichia_coli_Shigella/PDG00000004.1015/PDS000003441.82/70?key=C0CkKbge1dqijGXtyu15iOMmL04wcuZpItDvoYXB15ue2NaXlfba?treelabel=epi_type,isolation_source,biosample_acc)
- All (AMR): <https://www.ncbi.nlm.nih.gov/pathogens/isolates#/search/>
- Has phenotype:  
[https://www.ncbi.nlm.nih.gov/pathogens/isolates#/search/AST\\_phenotypes.\\*](https://www.ncbi.nlm.nih.gov/pathogens/isolates#/search/AST_phenotypes.*)
- AMR\_genotypes:blaKPC\* AND AST\_phenotypes:\*penem=S  
[https://www.ncbi.nlm.nih.gov/pathogens/isolates#/search/AMR\\_genotypes:blaKPC\\*%20AND%20AST\\_phenotypes:\\*penem=S](https://www.ncbi.nlm.nih.gov/pathogens/isolates#/search/AMR_genotypes:blaKPC*%20AND%20AST_phenotypes:*penem=S)

# Advanced searches

- 0-6 mindiff, submitted in September:

[https://www.ncbi.nlm.nih.gov/pathogens/isolates/#/search/mindiff:%5B0%20TO%206%5D%20%20AND%20creation\\_date:%5B2018-08%20TO%202018-09%5D?newtree=on](https://www.ncbi.nlm.nih.gov/pathogens/isolates/#/search/mindiff:%5B0%20TO%206%5D%20%20AND%20creation_date:%5B2018-08%20TO%202018-09%5D?newtree=on)

# Thank you.

Richa Agarwala  
Azat Badretdin  
Slava Brover  
Joshua Cherry  
Jinna Choi  
Vyacheslav Chetvernin  
Robert Cohen  
Michael DiCuccio  
Boris Fedorov  
Mike Feldgarden  
Lewis Geer  
Dan Haft  
Lianyi Han  
Avi Kimchi  
Michel Kimelman  
William Klimke  
Alex Kotliarov  
Valerii Lashmanov  
Aleksandr Morgulis  
Eyal Moses  
Chris O'Sullivan  
Arjun Prasad

Edward Rice  
Kirill Rotmistrovskyy  
Alejandro A. Schaffer  
Nadya Serova  
Stephen Sherry  
Sergey Shirayev  
Martin Shumway  
Oleg Shutov  
Alexandre Souvorov  
Tatiana Tatusova  
Igor Tolstoy  
Chunlin Xiao  
Leonid Zaslavsky  
Alexander Zasyarkin  
Lukas Wagner  
Hlavina Wratko  
Eugene Yaschenko  
  
David Lipman  
James Ostell  
Kim Pruitt

This research was supported by the  
Intramural Research Program of the  
NIH, National Library of Medicine.

[pd-help@ncbi.nlm.nih.gov](mailto:pd-help@ncbi.nlm.nih.gov)

CDC  
FDA/CFSAN  
GenFS  
USDA-FSIS  
PHE/FERA  
NIHGRI  
NIAID  
WRAIR  
Broad  
Wadsworth/MDH  
Vendors: PacBio, Illumina, Roche