MEDIA RELEASE
FOR IMMEDIATE RELEASE

22 February 2016

ACCELERATING GENOME ANALYSIS

Updated software SIFT 4G can further research in human health, the study of biological processes and agricultural products

SINGAPORE – An international team of scientists, led by researchers from A*STAR's Genome Institute of Singapore (GIS) and the Bioinformatics Institute (BII), have developed SIFT 4G[1] (SIFT for Genomes) – a software that can lead to faster genome analysis. This development was published in the scientific journal *Nature Protocols.*

Genome sequencing has been instrumental in improving knowledge of human diseases, by allowing scientists to understand their underlying biological mechanisms. It has also been critical in the global push towards precision medicine, where the genetic profile of a disease or patient could improve clinical decision-making in determining prognoses and the type of treatment prescribed for patients, paving the way for improved healthcare outcomes.

While technological advances have enabled the generation of vast amounts of data on the human body and other organisms, an issue facing the scientific community is the ability to analyse such great amounts of data well.

Based on the same principles as its predecessor, SIFT 4G can prepare predictions for genomes at a much higher speed. Once the predictions are ready, they are stored in a database and are ready for genome analysis. The use of graphics processing units (GPUs) instead of central processing units (CPUs) has resulted in much shorter prediction times and the ability to construct databases

---

[1] SIFT 4G is an updated version of SIFT (Sorting Intolerant from Tolerant), an algorithm that identifies which variants can cause disease or lead to different physical traits. SIFT has already been used in numerous disease, mutation, and genetic studies, and has been cited more than 6000 times according to Google Scholar.

for a large number of organisms. SIFT 4G already has predictions available for more than 200 organisms, thereby serving a larger research community.

While it previously took 25 days on 10 CPUs to create a database for SIFT predictions, it now takes just three days on a single GPU using SIFT 4G. By using a GPU instead of a CPU, the processing time on a protein decreased substantially from 4.2 minutes to 2.6 seconds. Once the predictions are ready, the researchers can use the database to analyse the genomes in just five minutes.

The increased efficiency in genome analysis will empower the research community in advancing science and developing technology that can benefit human health. The bioinformatics and genomic capabilities developed at BII and GIS enable both institutes to play a key role in quickly and accurately interpreting biological data to understand gene function, their interactions and the development of diseases.

Besides its obvious benefits to improving healthcare outcomes, genome sequencing has brought about significant advances in agricultural fields, as well as in basic research. Sequencing different breeds or strains of the same organism can lead to an understanding of the genetic basis of its observable characteristics.

For example, over 3,000 rice genomes were sequenced to interpret the genetic diversity that underlies traits such as cold tolerance and grain quality. These types of surveys can generate a huge impact, leading some to estimate the sequencing market in agricultural and other industrial applications to be valued at more than S$7 billion. Basic research also benefits from the rise of genome sequencing; for instance, the sequence of the *Drosophila*[2] genome has been applied to better understand the fundamentals of evolutionary processes.

Senior corresponding author of the research, GIS' Dr Pauline Ng – who created the SIFT algorithm with her PhD advisors, Prof Steven Henikoff and Jorja Henikoff, at the Fred Hutchinson Cancer Research Center (FHCRC) more than 15 years ago – is thrilled with what SIFT 4G can offer.

Dr Ng said, "I'm excited that SIFT 4G will make faster discovery possible for researchers studying mutations in their organism of choice. Any researcher who is studying genetic variation in an organism (whose genome has been sequenced) can now characterise their missense mutations with SIFT 4G. Its expanded functionality for many genomes will now enable researchers worldwide to use it as a resource across diverse projects."

---

[2] *Drosophilia* is a genus of small flies (often referred to as "fruit flies") often used as a model organism in developmental biology.

BII Executive Director Dr Frank Eisenhaber said, "Novel algorithms and processor architectures are essential for coping with the huge amounts of data collected in biological databases, which often outpaces the advances in computer performance. This is just one of the ways in which bioinformatics and computational biology can be utilised to advance research and understanding of biological processes."

GIS Executive Director Prof Ng Huck Hui said, "Time is of the essence, especially more so when it comes to the research fields. We have to be able to adapt quickly in order to keep up with the constantly evolving healthcare and biomedical landscape; SIFT 4G is a powerful tool for us to do so. It will accelerate the time taken for genome analysis and that can only benefit the research community and the public at large."

IMAGE
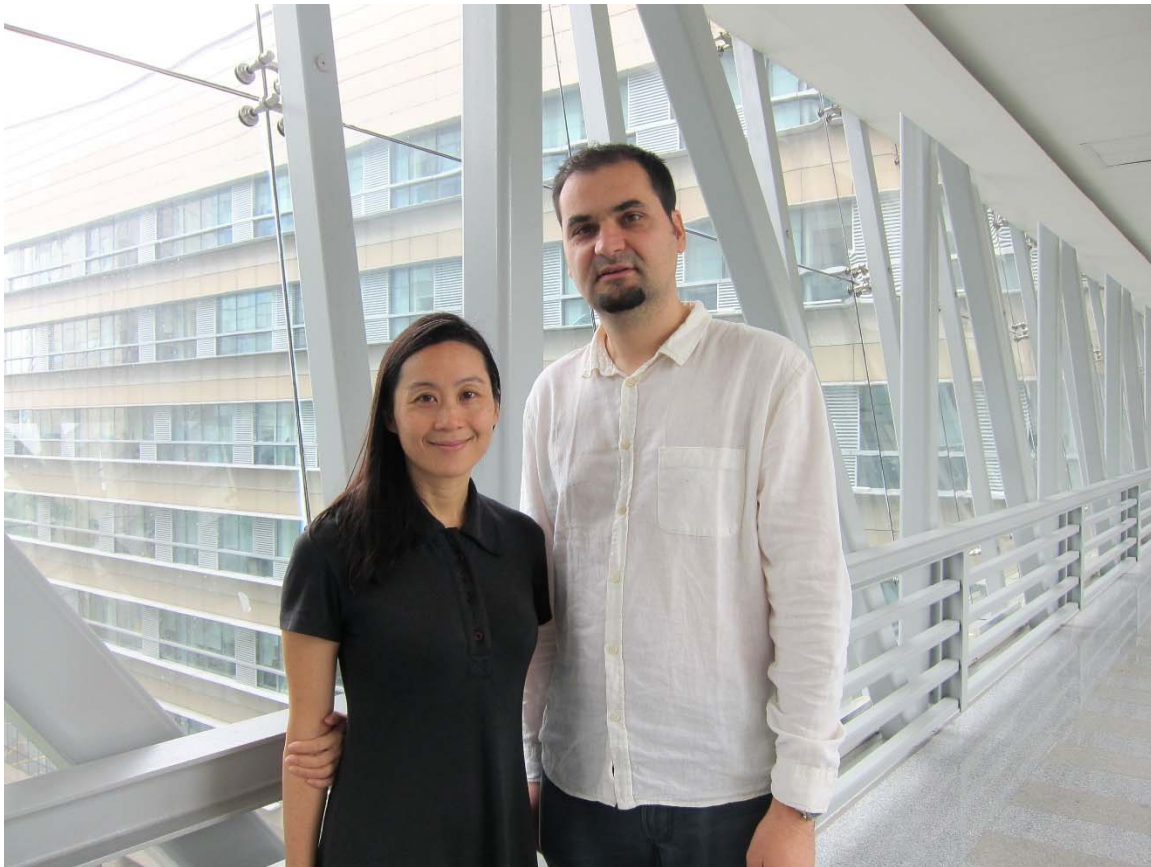
A*STAR researchers who developed SIFT 4G: Drs Pauline Ng (left) and Mile Sikic, from the Genome Institute of Singapore (GIS) and Bioinformatics Institute (BII)

Notes to Editor:

The research findings described in this media release can be found in the scientific journal *Nature Protocols,* under the title, "SIFT missense predictions for genomes" by Robert VASER[1,4], Swarnaseetha ADUSUMALLI[2,4], SIM Ngak Leng[2], Mile SIKIC[1,3] & Pauline C NG[2]

[1]Faculty of Electrical Engineering and Computing, University of Zagreb, Croatia
[2]Computational and Systems Biology Group, Genome Institute of Singapore, Singapore
[3]Bioinformatics Institute, Singapore
[4]These authors contributed equally to this work

*Senior Corresponding author: ngpc4@gis.a-star.edu.sg (P.C.N.)

Full text of the *Nature Protocols* paper can be accessed online from:
http://www.nature.com/nprot/journal/v11/n1/abs/nprot.2015.123.html

_____

For media queries and clarifications, please contact:

Joyce Ang
Senior Officer, Office of Corporate Communications
Genome Institute of Singapore, A*STAR
Tel: +65 6808 8101
Email: angjj@gis.a-star.edu.sg

Vanessa Loh
Senior Officer, Corporate Communications
Agency for Science, Technology and Research
Tel: +65 6826 6395
Email: vanessa_loh@a-star.edu.sg

_____

About A*STAR's Genome Institute of Singapore (GIS)

The Genome Institute of Singapore (GIS) is an institute of the Agency for Science, Technology and Research (A*STAR). It has a global vision that seeks to use genomic sciences to achieve extraordinary improvements in human health and public prosperity. Established in 2000 as a centre for genomic discovery, the GIS will pursue the integration of technology, genetics and biology towards academic, economic and societal impact.

The key research areas at the GIS include Human Genetics, Infectious Diseases, Cancer Therapeutics and Stratified Oncology, Stem Cell and Regenerative

Biology, Cancer Stem Cell Biology, Computational and Systems Biology, and Translational Research.

The genomics infrastructure at the GIS is utilised to train new scientific talent, to function as a bridge for academic and industrial research, and to explore scientific questions of high impact.

For more information about GIS, please visit www.gis.a-star.edu.sg.

About A*STAR's Bioinformatics Institute (BII)

The Bioinformatics Institute (BII) is an institute of the Agency for Science, Technology and Research (A*STAR) set up in July 2001. With a multi-disciplinary focus and collaborative outlook, BII recognises the need for depth and breadth in all its activities for building a thriving world-class biomedical research, graduate training and development hub in Singapore. In addition, BII is proactively involved in building a national resource centre in bioinformatics to meet the evolving needs of the scientific community in Singapore. The spectrum of research activities in BII includes theoretical approaches aimed at understanding biomolecular mechanisms that underlie biological phenomena, the development of computational methods to support this discovery process, and experimental verification of predicted molecular and cellular functions of genes and proteins with biochemical methods. BII also has a division of translational research aimed at enhancing applied research and industry collaborations..

For more information on BII, please visit: www.bii.a-star.edu.sg

About the Agency for Science, Technology and Research (A*STAR)

The Agency for Science, Technology and Research (A*STAR) is Singapore's lead public sector agency that spearheads economic oriented research to advance scientific discovery and develop innovative technology. Through open innovation, we collaborate with our partners in both the public and private sectors to benefit society.

As a Science and Technology Organisation, A*STAR bridges the gap between academia and industry. Our research creates economic growth and jobs for Singapore, and enhances lives by contributing to societal benefits such as improving outcomes in healthcare, urban living, and sustainability.

We play a key role in nurturing and developing a diversity of talent and leaders in our Agency and Research Institutes, the wider research community and industry. A*STAR oversees 18 biomedical sciences and physical sciences and engineering research entities primarily located in Biopolis and Fusionopolis.

For more information on A*STAR, please visit www.a-star.edu.sg.