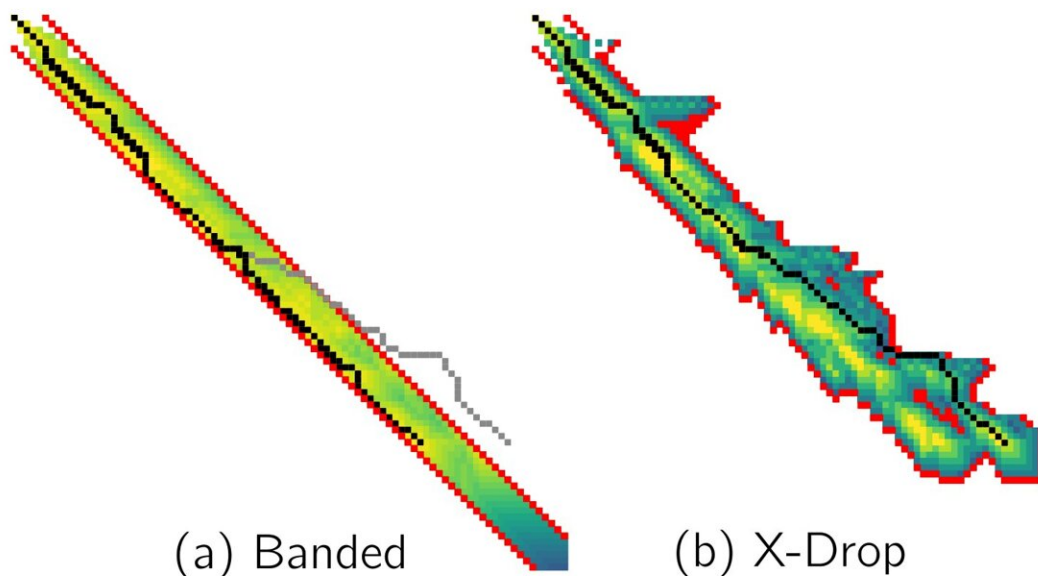


# Processor made for AI speeds up genome assembly

November 1 2023, by Patricia Waldron

---



On the left, the alignment is forced to stay within the banded area regardless of the score, missing the optimal alignment (gray). On the right, when the score (yellow-blue) goes below the current best score, the search is terminated (red boundary), and the optimal alignment (black) is returned. Credit: *arXiv* (2023). DOI: 10.48550/arxiv.2304.08662

A hardware accelerator initially developed for artificial intelligence operations successfully speeds up the alignment of protein and DNA molecules, making the process up to 10 times faster than state-of-the-art

methods.

This approach can make it more efficient to align [protein sequences](#) and DNA for [genome assembly](#), which is a [fundamental problem](#) in computational biology.

Giulia Guidi, assistant professor of computer science in the Cornell Ann S. Bowers College of Computing and Information Science, led a study to test the performance of the accelerator, called an intelligence processing unit (IPU), using existing DNA and protein sequence data. The IPU accelerates the alignment process by providing more memory to speed up data movement—a common holdup.

"Sequence alignment is an extremely important and compute-intensive part of basically any computational biology workload," Guidi said. "It is extremely common and it's usually one of the bottlenecks of the computation."

The study, "[Space Efficient Sequence Alignment for SRAM-Based Computing: X-Drop on the Graphcore IPU](#)," will be presented by co-first author Luk Burchard, a former visiting scholar at Cornell and doctoral student at Simula Research Laboratory, at the [Supercomputing2023 conference](#), Nov. 14. Max Xiaohang Zhao, also a former visiting scholar at Cornell, now at Charité Universitätsmedizin, is also a co-first author.

In her research, Guidi wants to help scientists solve problems they haven't even attempted yet because they require so much computational power. These [complex problems](#) require large-scale computation—assemblages of processors, memory, networks and [data storage](#) that can handle big computing tasks.

Aligning sequences of DNA or proteins is one of these complex

problems. When sequencing a genome, biologists end up with thousands or millions of short DNA sequences that must be put together like a puzzle. They use an algorithm to identify pairs of sequences that overlap, and then link up the pairs.

In the past decade, scientists have turned to graphics processing units (GPUs)—initially developed to accelerate graphics rendering in video games—to speed up sequence alignment by running calculations in parallel. With the development of IPUs for AI applications, Guidi and her colleagues wanted to know if they could harness the new accelerators to tackle this problem.

"The need for large-scale computation is growing for many domain sciences because we are so much better at generating data now than ever before," Guidi said. "Parallel computing moved from being a luxury to something that is non-negotiable."

IPUs attracted Guidi because they have substantial on-device bandwidth for transferring data and can handle uneven and unpredictable workloads. X-Drop, a popular algorithm for aligning sequences, has a very irregular computation pattern. When two sequences are a match, the algorithm requires a lot of computation to determine the right alignment, but when they don't match, the algorithm just stops. GPUs struggle with this kind of irregular computation, but the IPU excelled.

When Guidi's group assembled sequences from the model organisms *E. coli* and *C. elegans* with the help of the IPU, they achieved 10-times faster performance compared to a GPU, which spends too much time transferring data unnecessarily, and 4.65-times faster performance than a [central processing unit](#) (CPU) on a supercomputer.

Currently, what is limiting the size of the genomes scientists can process is the number of IPU and GPU devices available, as well as the

bandwidth for [data transfer](#) between the host CPU and the hardware accelerator. There is a lot of memory on the IPU, but transferring the data from the host causes a major bottleneck.

The team helped to address this issue by shrinking the memory footprint of the X-Drop algorithm by 55 times. This enabled it to run on the IPU and reduce the amount of data transferred from the CPU. As a result, the system could run larger comparisons and perform more of the sequence comparisons on the IPU, which helped to balance the uneven workload.

"You can exploit the IPU high memory bandwidth, which allows you to make the whole processing faster," Guidi said.

If vendors can upgrade the data transfer process between the CPU and IPU, and improve the software ecosystem, Guidi expects that she could process bigger genomes on the same IPUs.

"The IPU may become the next GPU," she said.

The study is [published](#) on the *arXiv* preprint server.

**More information:** Luk Burchard et al, Space Efficient Sequence Alignment for SRAM-Based Computing: X-Drop on the Graphcore IPU, *arXiv* (2023). [DOI: 10.48550/arxiv.2304.08662](https://doi.org/10.48550/arxiv.2304.08662)

Provided by Cornell University

Citation: Processor made for AI speeds up genome assembly (2023, November 1) retrieved 27 April 2024 from <https://phys.org/news/2023-11-processor-ai-genome.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private

study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.