

# GPU-BLASTN: Accelerating Nucleotide Sequence Alignment by GPUs

Kaiyong Zhao<sup>1</sup>, Xiaowen Chu<sup>1,2,\*</sup>

<sup>1</sup>: Department of Computer Science, Hong Kong Baptist University.

<sup>2</sup>: Institute of Computational and Theoretical Studies, Hong Kong Baptist University.

\*: To whom correspondence should be addressed.

Emails: KZ (kyzhao@comp.hkbu.edu.hk); XC (chxw@comp.hkbu.edu.hk)

The BLAST software package for sequence alignment is one of the most fundamental and widely used bioinformatics tools [1] [2]. Given the large population of BLAST users, any improvement in the execution speed of BLAST will bring significant benefits to the bioinformatics community. Some research groups have used GPUs to accelerate the speed of BLAST. E.g., GPU-BLAST uses GPUs to accelerate BLASTP, and it achieves 3 to 4 times of speedup over single-thread CPU based NCBI-BLASTP [3]. GPUs have also been successfully used to accelerate other sequence alignment tools, e.g., [4].

In this poster, we show our design, implementation, optimization, and experimental results of GPU-BLASTN, a GPU-accelerated version of the widely used NCBI-BLASTN. To the best of our knowledge, this is the first work that provides a complete solution for accelerating BLASTN by GPUs. GPU-BLASTN can obtain identical results as NCBI-BLASTN, and its speed on a contemporary Nvidia GTX680 GPU card is about 10 to 20 times faster than the speed of single-thread NCBI-BLASTN running on Xeon E5620.

We evaluate GPU-BLASTN by running sequence search experiments against human build 36 and mouse build 36 genome databases that have been masked with WindowMasker. We use six sets of query sequences with different lengths ranging from hundreds to hundreds of thousands. We compare the results and running time of GPU-BLASTN with those of NCBI-BLASTN on both single-thread CPU and multi-thread CPU.

The GPU-BLASTN will be open source and freely available to the bioinformatics community.

## References

1. Altschul,S.F. et al. (1990) Basic local alignment search tool. *J. Mol. Biol.*, 215, 403–410.
2. Altschul,S.F. et al. (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.*, 25, 3389–3402.
3. P. D. Vouzis and N. V. Sahinidis (2011) GPU-BLAST: using graphics processors to accelerate protein sequence alignment. *Bioinformatics*, 27(2), 182-188.
4. C. M. Liu et. al. (2012) SOAP3: ultra-fast GPU-based parallel alignment tool for short reads. *Bioinformatics*, 28(6), 878-879.