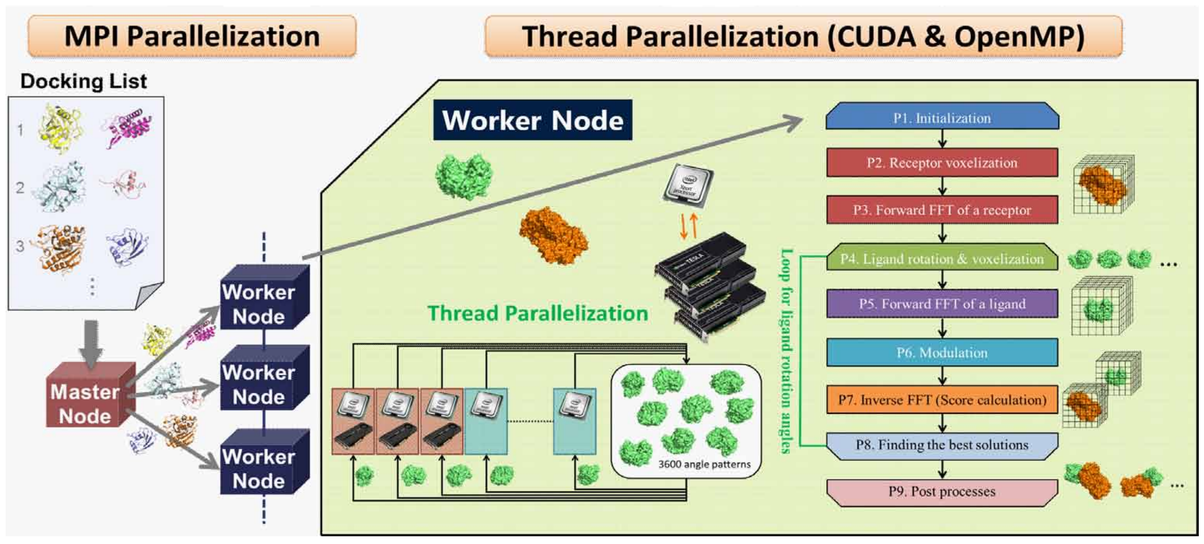
Supercomputing for Bioinformatics

GPU-acceleration on TSUBAME 2.5

TOKYO TECH Pursuing Excellence

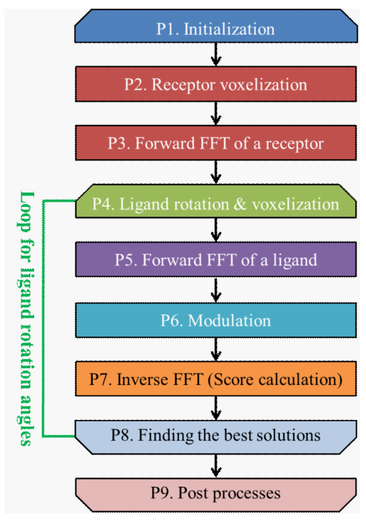
MEGADOCK-GPU: GPU-acceleration of protein-protein docking

MEGADOCK [Ohue *et al.*, *Bioinformatics,* (2014)]



- Protein docking (rigid-docking based):
  - original rPSC model + electrostatic interaction
  - 3-D convolution by FFT: $O(n^6) \rightarrow O(n^3 \log n)$
- Protein-Protein Interaction (PPI) prediction:
  - post-docking analysis by clustering and threshold logic
  - OpenMP/MPI hybrid parallelization for all-to-all analysis
- PPI network prediction:
  - application to bacterial chemotaxis (about 100x100 pairs)
  - application to lung cancer pathway (about 500x500 pairs)

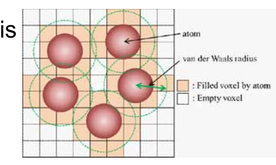
GPU implementation [Shimoda *et al.*, *ParBio2013,* (2013)]



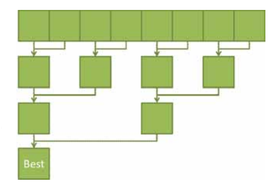
P5, P7: Forward FFT & Inverse FFT
This process is accelerated using CUFFT library.

P6: Modulation
This process performs complex conjugates and multiplication of the results of forward FFT, and is parallelized by element

$$\mathrm{FFT[R]}^* \times \mathrm{FFT[L]}$$

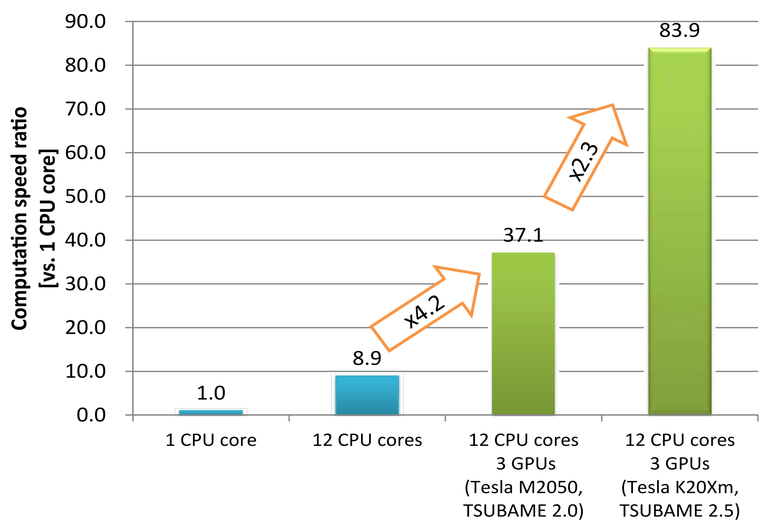


P4: Ligand voxelization
This process assigns a value to each grid within van der Waals radius of the atom, and parallelized for each atom



P8: Finding the best solutions
This process selects the best docking poses according to the docking score, and accelerated using reduction method

Speedup by GPGPU



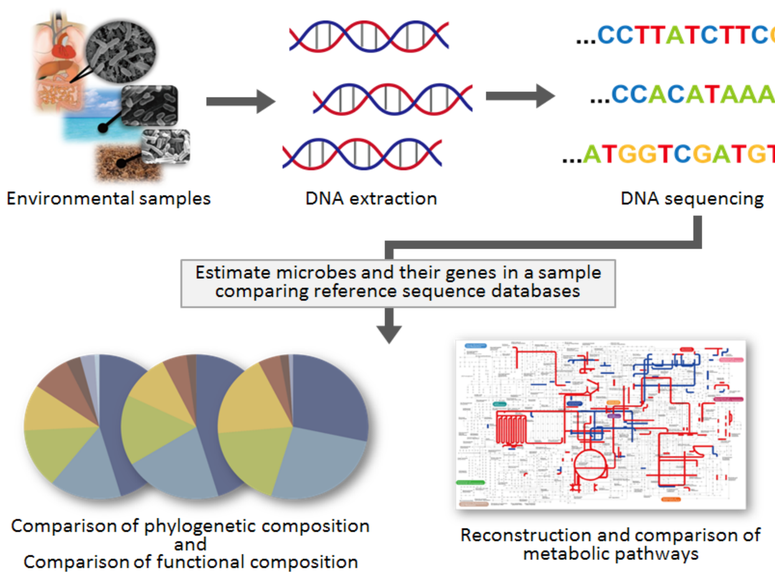
Docking calculations for 352 protein pairs

- 83.9-fold speedup with 3 GPUs on TSUBAME 2.5

GHOSTM: GPU-acceleration of sequence homology search

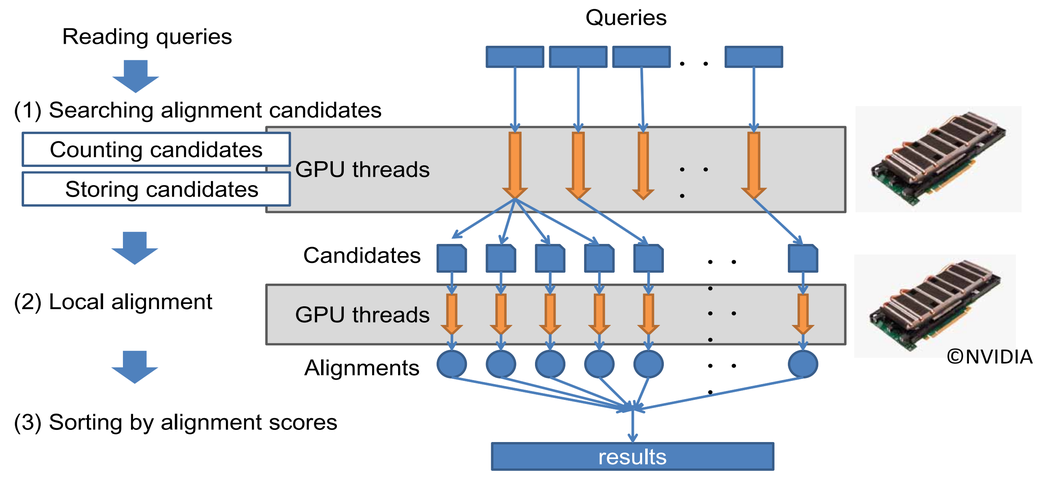
- Fast sequence homology search for metagenomics

  [Suzuki *et al.*, *PLoS ONE*, (2012)]
  - enough sensitive for metagenomic annotation
  - similar to BLAST but optimized for GPU-calculation
  - implemented on GPUs by NVIDIA CUDA
  - approximately 130 times faster than NCBI BLAST

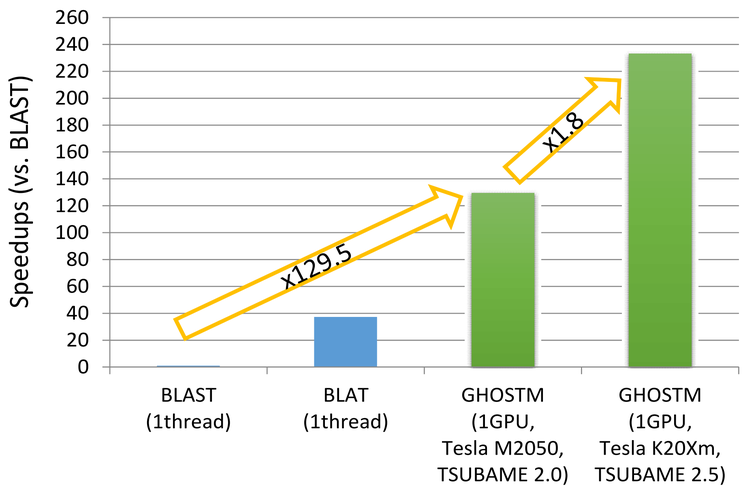


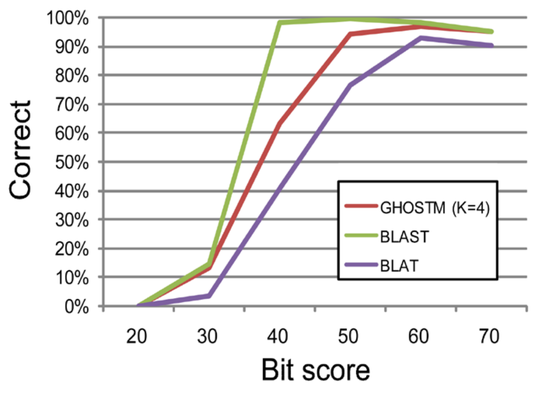
Metagenome analysis: the study of the genomes of uncultured microbes obtained directly from microbial communities in their natural habitats such as soils, seas, and human bodies.

Data flow and processing



Search speed and sensitivity





Correct answer: Smith-waterman local alignment algorithm by using SSEARCH

- 233-fold speedup with 1 GPU on TSUBAME 2.5

http://www.gsic.titech.ac.jp/sc14