

## SC06 Schedule

### High-Performance Computing Methods for Computational Genomics

**Session:** M11

**Event Type:** Tutorial

**Date:** Monday, Nov 13

**Time:** 1:30pm - 5:00pm

**Presenter(s):** Srinivas Aluru, David Bader, Ananth Kalyanaraman

**Location:** 17

#### Abstract:

The high computational requirements of several applications in computational genomics are aggravated by an exponential growth in biological databases. This tutorial will provide a detailed introduction to high-performance computing methods designed to address various large-scale problems in computational genomics. First, we will describe mpiBLAST and ScalaBLAST, which are parallelizations of the NCBI BLAST suite of programs used for querying against large sequence databases. Next, we will describe PaCE, which is a parallel DNA sequence clustering algorithm with applications to clustering Expressed Sequence Tags and whole genome assembly. Next, we describe GRAPPA, which is a high-performance software suite developed for phylogenetic reconstruction of a collection of organisms or genes. Throughout the tutorial, emphasis will be on scalability and effectiveness in exploiting large-scale state-of-the-art supercomputing technologies.

The intended audience are academic and industry researchers, educators, and/or commercial application developers, with a computational background. No background in biology is assumed.

Awards: Best Paper Nomination

Introductory: 40% Intermediate: 40% Advanced: 20%