

Computational infrastructure  
Division of Bioinformatics  
Medical University of Innsbruck

# Division of Bioinformatics

## Medical University of Innsbruck

- Personnel resources
  - 1 Professor, 3 assistant professors
  - 3 Postdocs, 3 PhD students
- Research activities
  - Computational genomics
  - Cancer immunology

# Current HPC HW infrastructure

- 6-node HPC cluster
  - 1 x Master node + 5 x Slave nodes
    - 288 Intel Xeon cores
    - 3 TB RAM (512 GB / node)
    - 1792 CUDA GPU cores / 18 GB RAM
    - 1 TB local scratch disk / node
- 456 TB SAN storage
  - 2 FC attached NFS servers
  - 11 Volumes (9 - 60TB each)

# Software in use

- OpenSource bioinformatics tools
  - NGS data analysis
  - R, R Bioconductor
  - Perl, Python, Java, C, C++
  - Currently > 130 software tools installed
    - Many tools support SMP, but few MPI or similar
      - computing nodes with many cores instead of many nodes with few cores
- Matlab
  - distributed Computing Server
  - Parallel Computing Toolbox

# Scientific use cases

- Cancer immunology:
  - Analyses of genomics data (e.g. TCGA) with respect to immunological parameters (Angelova et al., Genome Biology 2015)
  - Development of analytical tools for cancer immunotherapy
    - Horizon 2020 project APERIM (coordinator Z. Trajanoski):  
**A**dvanced Bioinformatics Tools for **PER**sonalised Cancer **IM**munotherapy

# Bottlenecks/Problems

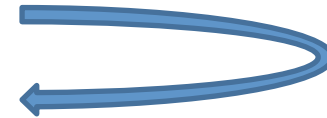
- Storage space

- TCGA data ~2.5 PB and growing <-> ~500 TB local

- Download → analyze → discard → new?

- ...weeks...

- ...weeks...



- Storage performance

- Multiple compute engines/cores accessing same data/volume → high IO-load on NFS