

The Lifelines high performance computing (HPC) cluster

Via the Lifelines high performance computing (HPC) environment researchers are able to analyze the genetic data from Lifelines participants. Access to the genetic data is provided to researchers that have an approved proposal in which they describe their research and elaborate on the use of the requested data. If the proposal is approved, researchers will be able to access the HPC cluster via the secure virtual working environment called 'workspace', using either the terminal PuTTY or WinSCP. The data managers from Lifelines will prepare and configure these applications for use, allowing easy access.

Linux environment

The HPC cluster is a Linux environment and requires command-line commands for interaction. Therefore, working with the genetic data requires knowledge on using Linux and understanding programming language. If researchers are unfamiliar with this approach, it is highly recommended to include a member to the project team that has a background in bio-informatics or has experience in working with similar clusters.

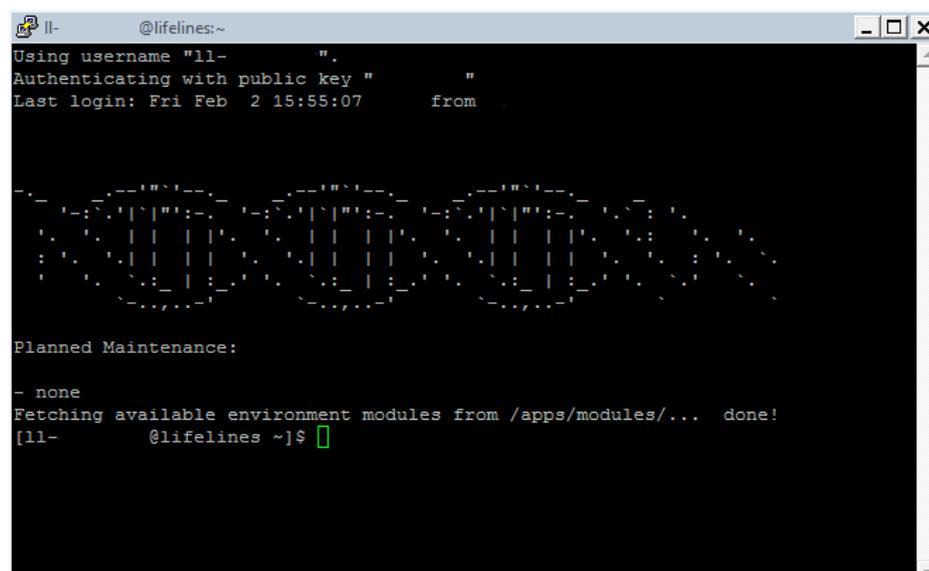
The cluster provides a wide variety of software to assist in genetic analysis as well as more general scripting. If researchers require additional software, a request can be send to Lifelines data management (data@lifelines.nl). To execute a script or an analysis, a job can be submitted which will be scheduled by the SLURM workload manager. Subsequently, the status of the job can be monitored on its progress. Via WinSCP, files can easily be transferred between the workspace and folders on the cluster.

More information

If you would like to learn more about our HPC cluster and its services, feel free to visit the following website: http://wiki.gcc.rug.nl/wiki/HPC_analysis. To gain access to the Lifelines directories, a relevant proposal including a catalogue selection should be submitted.

Genotypic data

The genome-wide genotype data is based on the Illumina CytoSNP-12v2 array and is currently available for ~15,000 samples. All independent and Caucasian-ancestry samples have also been imputed using the Genome of the Netherlands (GoNL) release 5 and the 1000 Genomes phase1 v3 reference panels.



```

@lifelines:~
Using username "ll-".
Authenticating with public key " "
Last login: Fri Feb  2 15:55:07    from

Planned Maintenance:

- none
Fetching available environment modules from /apps/modules/... done!
[ll- @lifelines ~]$
  
```