# DOWNLOADING OF NECESSARY DATA AND SOFTWARE.

Check if java is installed. If not please download it over here:

http://java.com/download/index.jsp

Download the necessary software and data.

Software: http://129.125.135.180/Exchange/eQTL-Mapping/GenotypeAndExpressionData.zip

Data: http://129.125.135.180/Exchange/eQTL-Mapping/Software.zip

# Optional

If you are working on windows xp and have administer rights please also install powershell.

http://www.microsoft.com/en-us/download/details.aspx?id=7217

Powershell is not necessary but its more user friendly than using the command prompt. Windows 7 and newer have powershell installed by default.

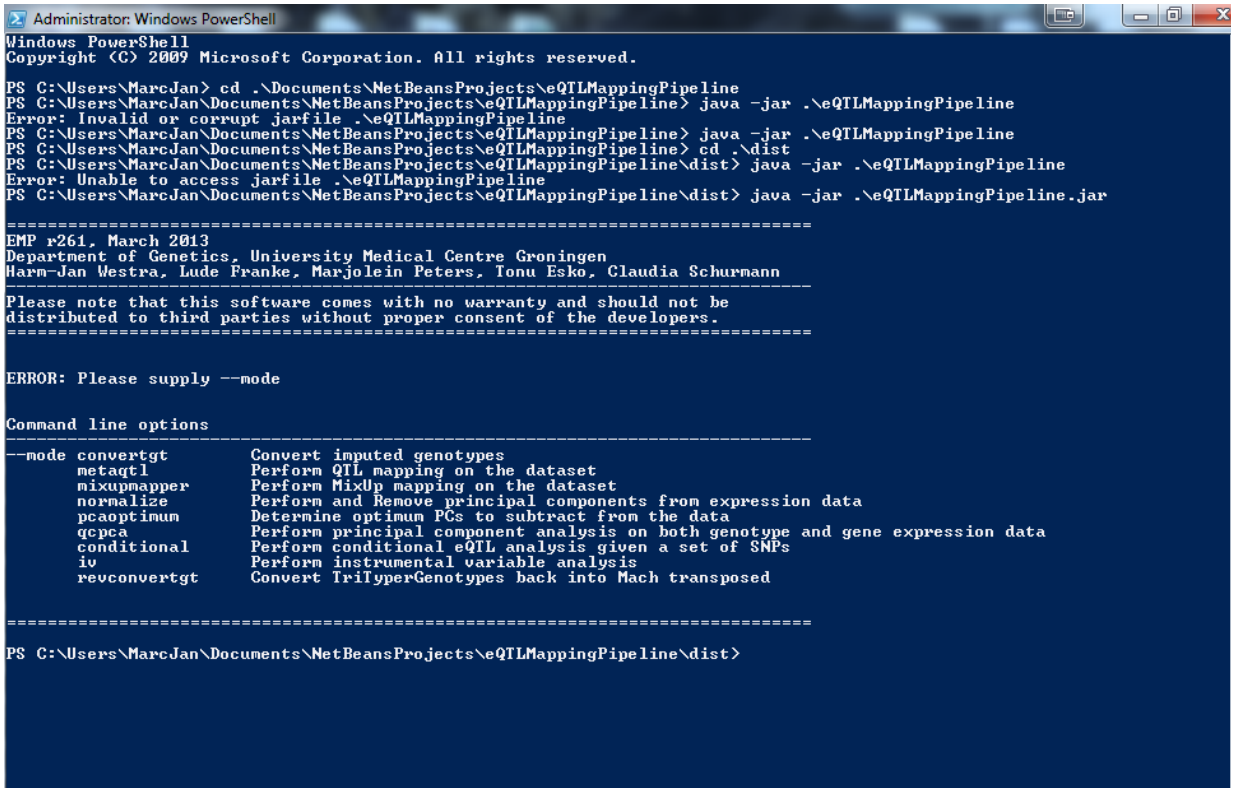
# TEST THE PROGRAM

Open the terminal, go to the folder where you downloaded the software.

Use "`cd ./destination-folder`" to do so. When you are in the correct folder type:

"`java -jar ./Software/eQTLMappingPipeline.jar `"

Please check if you get this output:



If so the software is ready to run. All of the steps below shouldn't take more than 15 minutes (run-time). If you are waiting for more than 15 minutes please ask for help!

# PRETREATMENT OF THE DATA

Before we can start with an eQTL mapping we first have to normalize the data. We start off with raw expression data, supplied in the data zip. As a first step we are going to use a quantile normalization, to make the samples comparable to each other. During the quantile normalization the distribution of expression values is made identical for all samples. After this we do a log transformation of the expression values. Followed by standardizing the probe values across all samples, to do so we set the mean of a probe to 0. As a last step we also do this per sample, so after this the standard deviation per samples 1 and mean is 0.

Use: "`java –jar ./Software/eQTLMappingPipeline.jar --mode normalize`" to do so.

Specify the input, output and the steps described in the help text when running the command above. Otherwise you will have to wait a long time!

# RUNNING YOUR FIRST EQTL MAPPING.

After the normalization you can run your first eQTL analysis. See if you can find eQTLs using this command: "`java –jar ./Software/eQTLMappingPipeline.jar --mode metaqtl`". Again check the help text when running the command to see what input is necessary. You need to specify "`—in`", "`—cis`" "`—perm`", "`—inexp`", "`—gte`", "`—inexpannot`" files to do the mapping.

Check the P-values of the top eQTLs and the total number of eQTLs in the file: "eQTLsFDR0.05.txt". 10 permutations is enough for these analyses (`--perm 10`).

**Significance of the top P-value is?**
**Number of eQTLs is?**

Check in the plots directory for the plots of the top eQTL effects in the run. In the plots you can see the effect of a SNP op an expression probe.

# CHECK IF THERE ARE SAMPLE MIX-UPS.

During sample preparation in the lab or mislabeling we can create mixups in the data. This will have a negative influence on the ability to detect eQTLS. The software includes a piece that can detect mixups. Use: "`java –jar ./Software/eQTLMappingPipeline.jar --mode mixupmapper`" to find mixups in the data.

Check in the output of mixupmapper: "mixupmapper/BestMatchPerTrait.txt". Using this file you have to change your genotype to expression coupling file. Replace swaps and remove samples which are a best match to another expression / genotype sample.

# CHECK THE INFLUENCE OF SAMPLE MIX-UPS ON THE QTL ANALYSIS.

Do a QTL analysis on this cleaned data and check the influence of the mix-ups on the P-values of the QTLs and the number of QTLs.
Again use: "`java –jar ./Software/eQTLMappingPipeline.jar --mode metaqtl`" to do so. Keep in mind to use the changed –gte file.

**Significance of the top P-value is?**
**Number of eQTLs is?**

# CHECK THE INFLUENCE OF CONFOUNDER CORRECTION.

Use the software to do an extra normalization step. We want to do a principal component correction, we correct the data for confounders with the first 5 principal components. To do so use:

"`java –jar ./Software/eQTLMappingPipeline.jar`"

`Use these flags: --adjustPCA --maxnrpcaremoved 5 --stepsizepcaremoval 5`

After this do an eQTL mapping on the confounder and mix-up cleaned data. Check if what happened to the P-values and the number of QTLs.

**Significance of the top P-value is?**
**Number of eQTLs is?**

# INTERPRETATION OF THE QTLS

Select all the genes that are associated to an QTL and put them in this tool:
http://genenetwork.nl:8080/GeneNetwork/pathway.html

Check what the highest associated tissue corresponding to the genes are (displayed in the bottom) and what the functions of these genes are, this is displayed in the other part of the page.

One of the SNPs with an eQTLs you find is: rs10518693. This SNP is associated to Metabolic traits (http://www.ncbi.nlm.nih.gov/pubmed/21886157).

**Look up the SNP and find the gene the SNP effects. Find the function of the gene and check the position of the gene and SNP.**