Hands-on section #3: Cloud Workstation and Command Line Interface

Outline

1. Logging into DNAnexus via the Command Line Interface (5 minutes)
2. Command line help (2 minutes)
3. Navigate directories, make output directory, examine files (3 minutes)
4. Run single variant analysis for inverse-normal transformed LDL on chromosome 19 (10 minutes)
5. Run single variant analysis for inverse-normal transformed LDL on chromosome 19 specifying parameters on the command line (5 minutes)
6. Spend 15 minutes running another app from the command line (15 minutes)
7. Using the cloud workstation to work with data (15 minutes)

References:

- Command Line Interface https://wiki.dnanexus.com/Command-Line-Client/Quickstart
- Index of dx commands https://wiki.dnanexus.com/Command-Line-Client/Index%20of%20dx%20Commands

Task 1) <u>Logging into DNAnexus via the Command Line Interface</u>

- Open terminal and change to the directory where you have installed the SDK
- Load environment

```
$ source dx-toolkit/environment
$ dx login
```

Task 2) <u>Command line help</u>

```
$ dx help
```

Task 3) <u>Navigate directories, make output directory, examine files</u>

- **File paths:** <project>:/path/to/file.txt
- Example: workshop:/phenotypes/ldl_pheno.csv

List directory contents:

```
$ dx ls
$ dx ls /tools
$ dx ls workshop:/tools
```

List current directory:

```
$ dx pwd
```

Change directories:

```
$ dx cd output
$ dx cd /output
$ dx cd workshop:/output
```

Make directory:

```
$ dx mkdir YOURFILENAME
$ dx mkdir workshop:/tools/YOURFILENAME
```

Look at the first few lines of a file:

```
$ dx head /phenotypes/ldl_pheno.csv
$ dx head workshop:/phenotypes/ldl_pheno_invnt.csv
```

Task 4) Run single variant analysis for inverse-normal transformed LDL on chromosome 19.

Getting help on running apps

```
$ dx help exec
```

Examine app and get list of input parameters:

```
$ dx run workshop:/tools/genesis_v0.6 -h
$ dx describe workshop:/tools/genesis_v0.6
```

Run app and manually enter input parameters:

```
$ dx run workshop:/tools/genesis_v0.6
```

App parameters:

```
phenofile: workshop:/phenotypes/ldl_pheno_invnt.csv
```

```
outcome_name: LDL_adj
```

```
snpinfofile:
workshop:/annotation/for_analysis/WGS_Freeze3_UCSCAnnovar_SNPlist_lipidAnnot_chr19.txt
```

```
genotypefile:
workshop:/genotypes/GDS/CHGS_WGS_frz3_pld_chr19.postqc.recode.gds
```

```
outputfilename: ldl_chr19

kinshipmatrix: workshop:/kinship/WGS_pedigree_kinship.Rda

test_type: Single
```

Optional parameters:

```
(0) covariate_list: pc1,pc2
(3) pheno_id: idno
```

Task 5) Run single variant analysis for inverse-normal transformed LDL on chromosome 19 specifying parameters on the command line.

```
$ dx run workshop:/tools/genesis_v0.6 -iphenofile=workshop:/phenotypes/ldl_pheno_invnt.csv -ioutcome_name="LDL_adj" -isnpinfofile=workshop:/annotation/for_analysis/WGS_Freeze3_UCSCAnnovar_SNPlist_lipidAnnot_chr19.txt -igenotypefile=workshop:/genotypes/GDS/CHGS_WGS_frz3_pld_chr19.postqc.recode.gds -ioutputfilename=ldl_chr19 -ikinshipmatrix=workshop:/kinship/WGS_pedigree_kinship.Rda -itest_type=Single -icovariate_list="pc1,pc2" -ipheno_id="idno"
```

Task 6) Spend 15 minutes running another app from the command line.

- Suggestions:
  - qqplot app
  - mhplot app
  - genesis_single_QQ workflow

Task 7) Using the cloud workstation to work with data

```
$ dx describe cloud_workstation
```

```
$ dx run cloud_workstation --ssh
```

Parameters:

```
phenofile: workshop:/phenotypes/ldl_pheno_invnt.csv
```

Inside cloud workstation (can run typical linux commands)

```
》 ls -l ldl_pheno_invnt.csv
》R
```

in R:

```
in.pheno <- read.csv("ldl_pheno_invnt.csv")
summary(in.pheno[,1:8])
```

```
install.packages("sm")
library(sm)
pdf("ldl_by_cohort.pdf")
sm.density.compare(in.pheno$LDL_adj, in.pheno$study)
legend("topright", legend=levels(in.pheno$study), fill=2+(0:nlevels(in.pheno$study)))
dev.off()
```

```
》dx upload –h
》dx upload --destination workshop:/output/alisamanning/ldl_by_cohort.pdf ldl_by_cohort.pdf
》exit
```