

Exercises 1: expression and genotype data; annotation

VJ Carey

February 19, 2008

1. Attach the MAQCsubset package and bring afxsubRMAES into scope. Execute and interpret the following:

```
> plot(exprs(afxsubRMAES)[21087, ] ~ afxsubRMAES$pctBrain)
```

2. Interpret:

```
> t.test( exprs(afxsubRMAES[21087, afxsubRMAES$pctB==.75]),  
+        exprs(afxsubRMAES[21087, afxsubRMAES$pctB==1]) )
```

3. Attach the GGtools package and bring chr20GGdem into scope.

```
> library(GGtools)  
> data(chr20GGdem)
```

What is the meaning of the result of `annotation(chr20GGdem)`?

4. Execute the following, then decode the result into HUGO gene symbols:

```
> featureNames(chr20GGdem)[2000:2004]
```

On which chromosome do we find gene USP20 (use Bioconductor resources only)?

5. Interpret:

```
> table(snps(chr20GGdem)[2006, ])  
> table(snps(chr20GGdem)[2006, ], snps(chr20GGdem)[2007, ])
```

What are the dbSNP identifiers for these SNP? Use chr20meta in the GGtools package to determine their locations. Check the result with biomaRt.

6. Execute and comment in relation to the ‘rare allele count’ definition underlying the genotype data in chr20GGdem.

```
> snps(chr20GGdem)["rs167391", ]
```

Visualize, and discuss the relationship of PITPNC1 expression to genotype at rs167391. Formalize a hypothesis and test it.