HIGH FAT DIET INDUCES SEX SPECIFIC DIFFERENTIAL GENE EXPRESSION

Júlia Jiménez, Ainhoa López , Alba Peña

OVERVIEW

INTRODUCTION TO THE DATA

Structure of the Dataset: Metadata and Gene Expression Matrix

**SAMPLE METADATA :
BIOLOGICAL &
EXPERIMENTAL DESIGN**

```
> head(metadata)
             sex     Diet specimen_part Replicate
SRR8270499 female HighFat         head        R1
SRR8270501 female HighFat         head        R1
SRR8270495 female HighFat         head        R1
SRR8270500 female HighFat         head        R1
SRR8270496 female HighFat         head        R1
SRR8270497 female HighFat         head        R1
```

**GENE EXPRESSION MATRIX :
SAMPLES AND GENES**

```
> head(rownames(gene_data))
[1] "SRR8270499" "SRR8270501" "SRR8270495" "SRR8270500" "SRR8270496" "SRR8270497"
> head(colnames(gene_data))
[1] "FBgn0000003" "FBgn0000008" "FBgn0000014" "FBgn0000015" "FBgn0000017" "FBgn0000018"
```

PCA PLOTS:
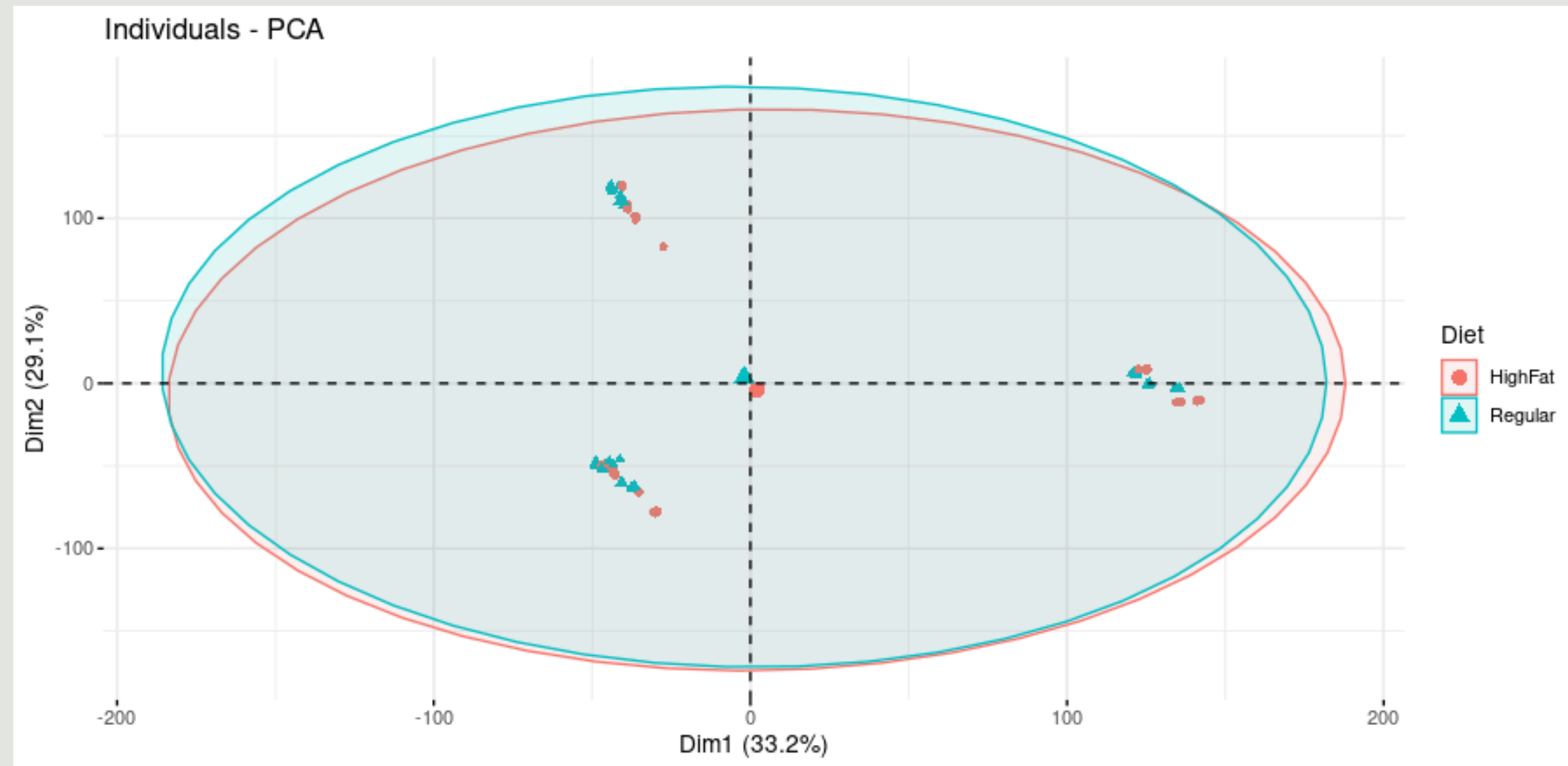
SIGNIFICANT FACTORS
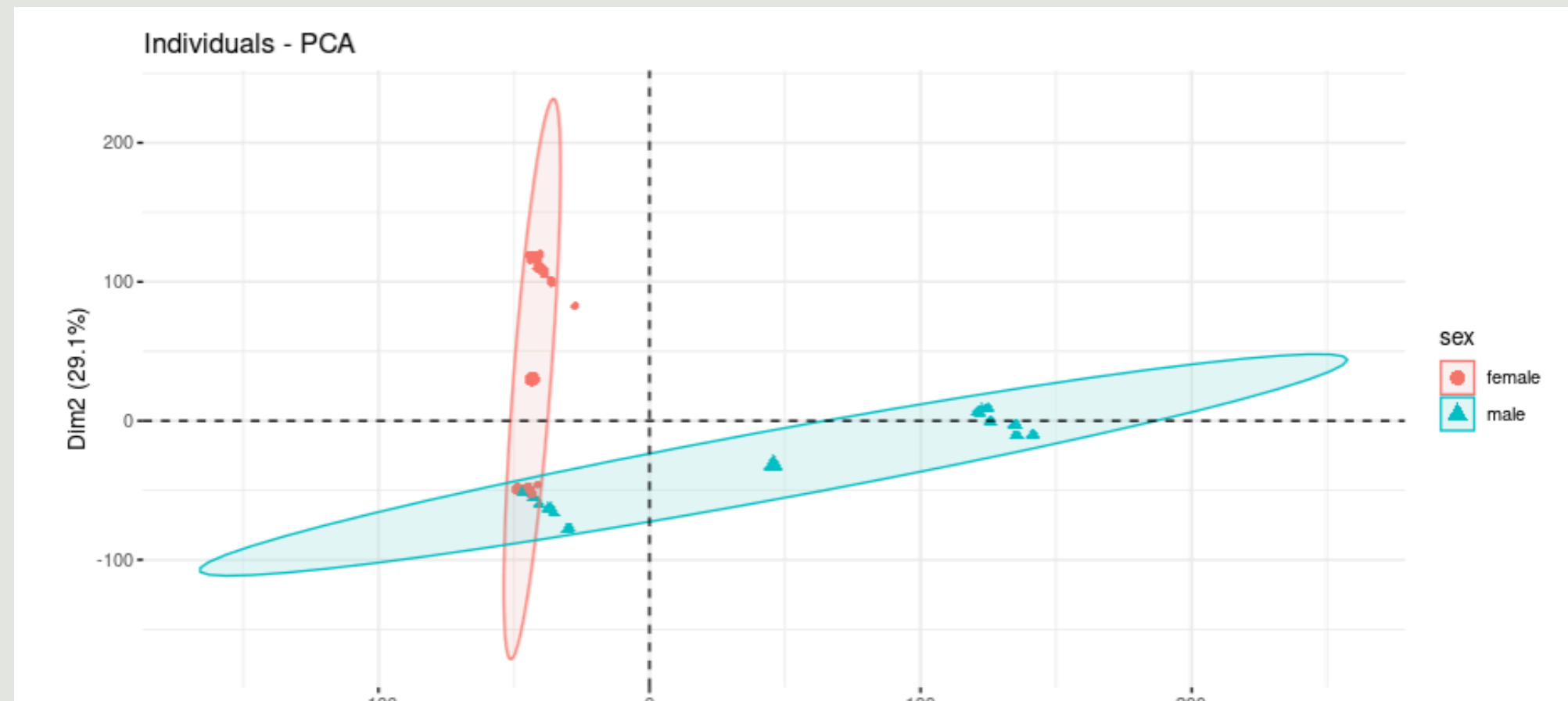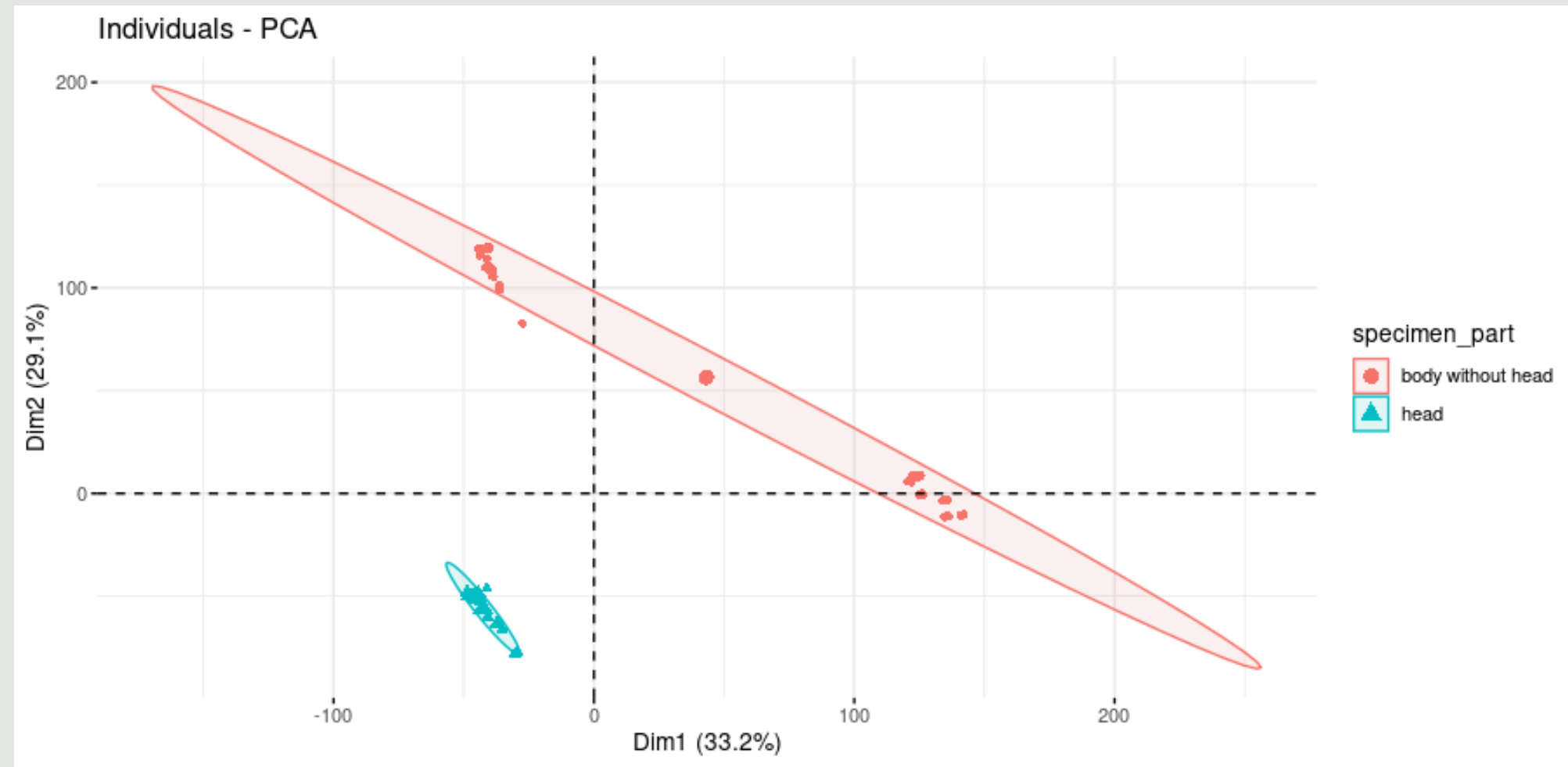
ANOVA TEST: SIGNIFICANT FACTORS

```
## Analysis of Variance Table
##
## Response: PC1
##                 Df Sum Sq Mean Sq F value    Pr(>F)
## sex             1 357784  357784  101.54 < 2.2e-16 ***
## Residuals     179 630741    3524
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' '
1
```

```
## Analysis of Variance Table
##
## Response: PC1
##                   Df Sum Sq Mean Sq F value    Pr(>F)
## specimen_part     1 333620  333620  91.186 < 2.2e-16 ***
## Residuals       179 654904    3659
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' '
1
```

PCA PLOT:

NOT SIGNIFICANT FACTOR

Individuals - PCA

ANOVA TEST:

NOT SIGNIFICANT FACTOR

```
## Analysis of Variance Table
##
## Response: PC1
##              Df  Sum Sq  Mean Sq  F value  Pr(>F)
## Diet          1     730    730.5   0.1324  0.7164
## Residuals   179  987794   5518.4
```
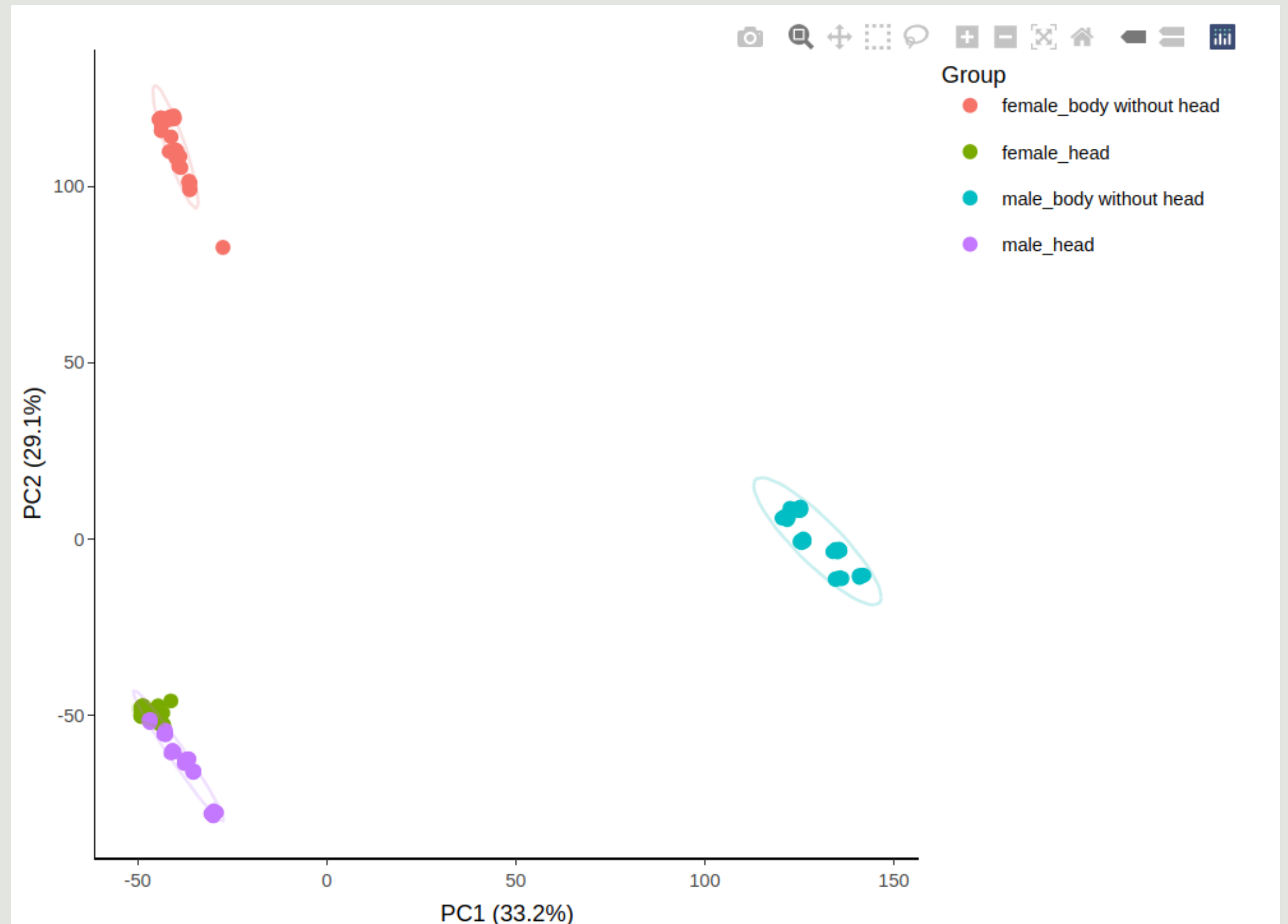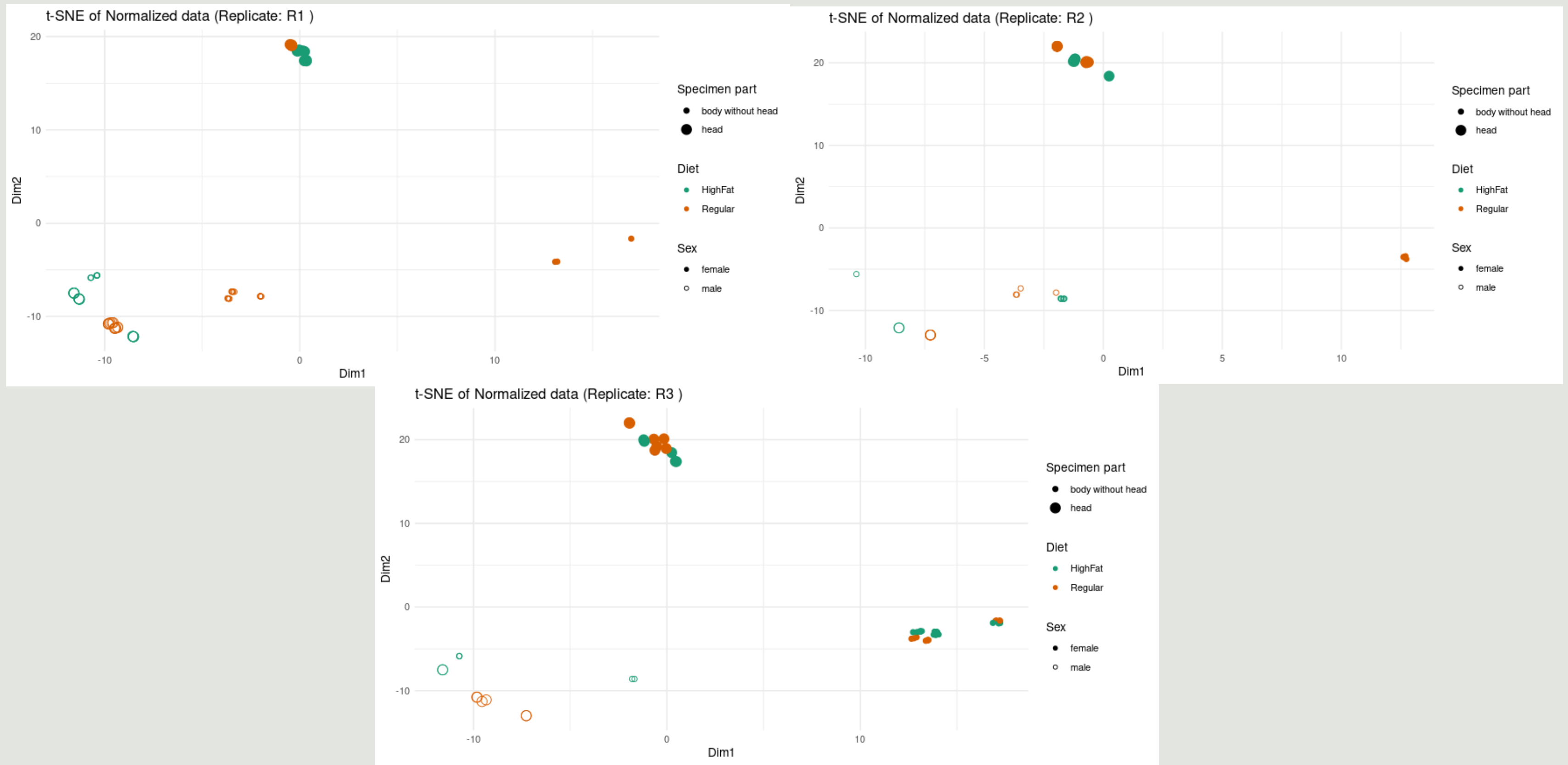
PCA : SEX AND BODY PART

The distinct clustering of the four groups shows that gene expression is strongly influenced by both sex and tissue. There is clear separation across both PC1 and PC2, confirming that these two variables interact and contribute significantly to transcriptomic variation.
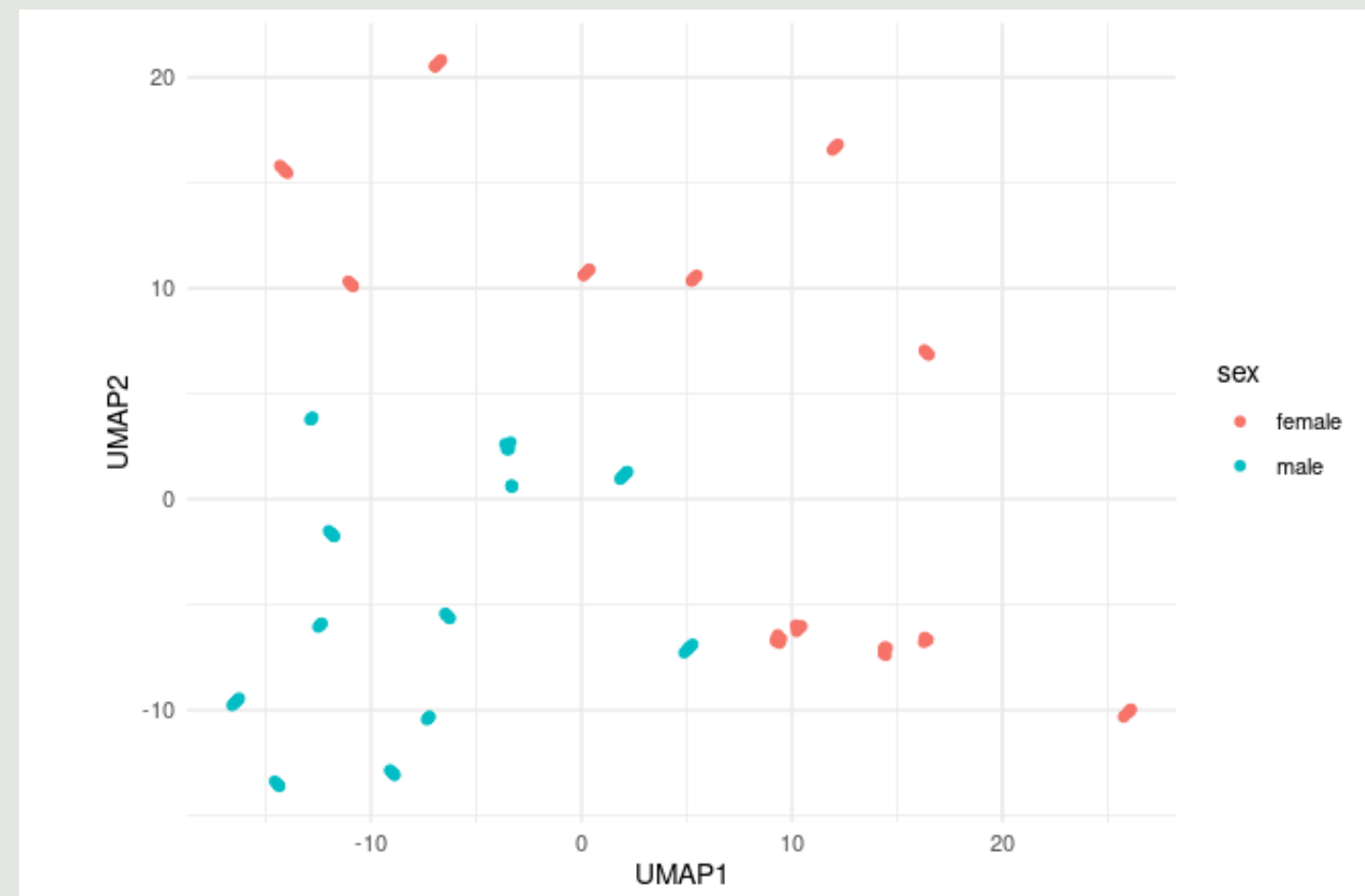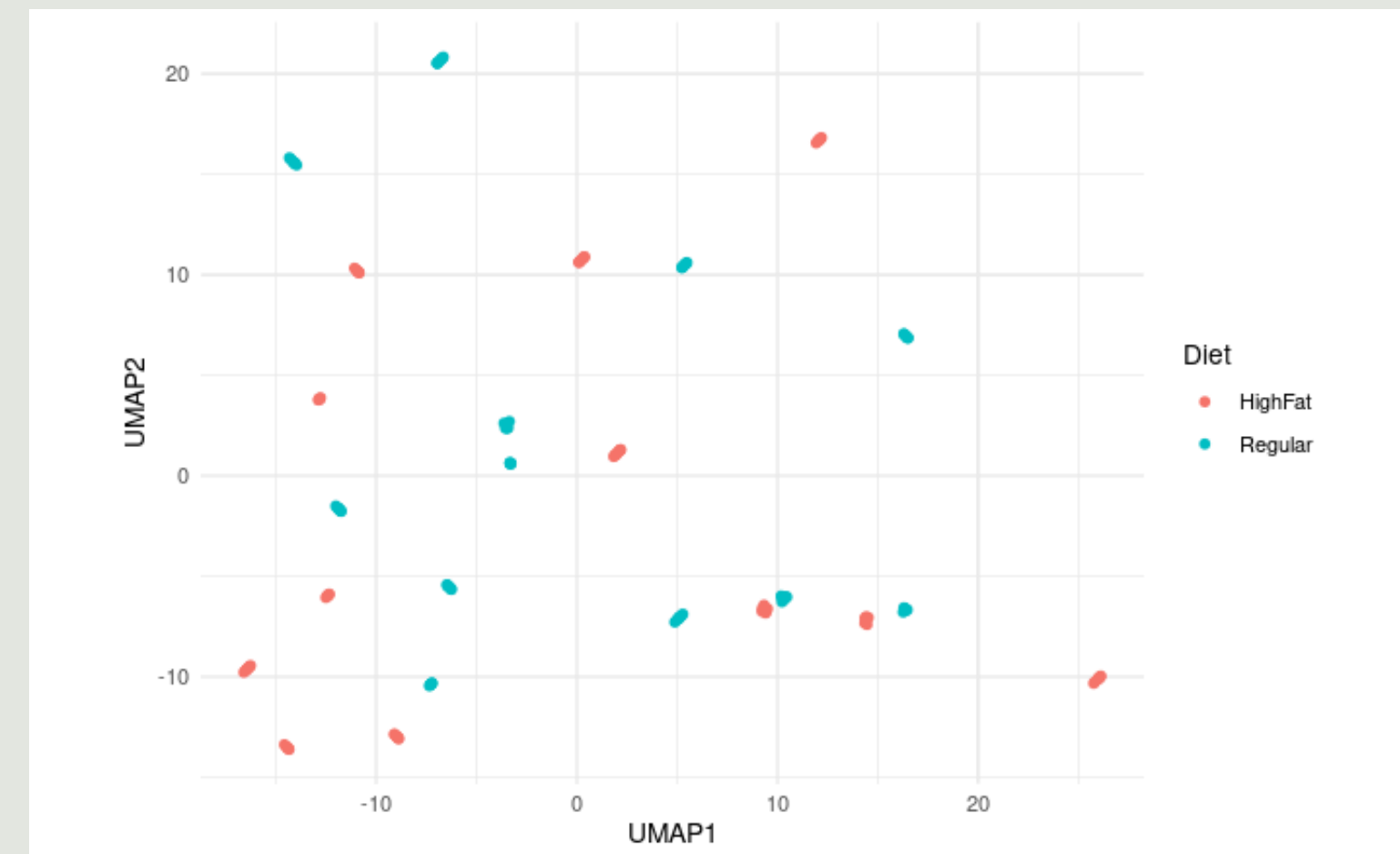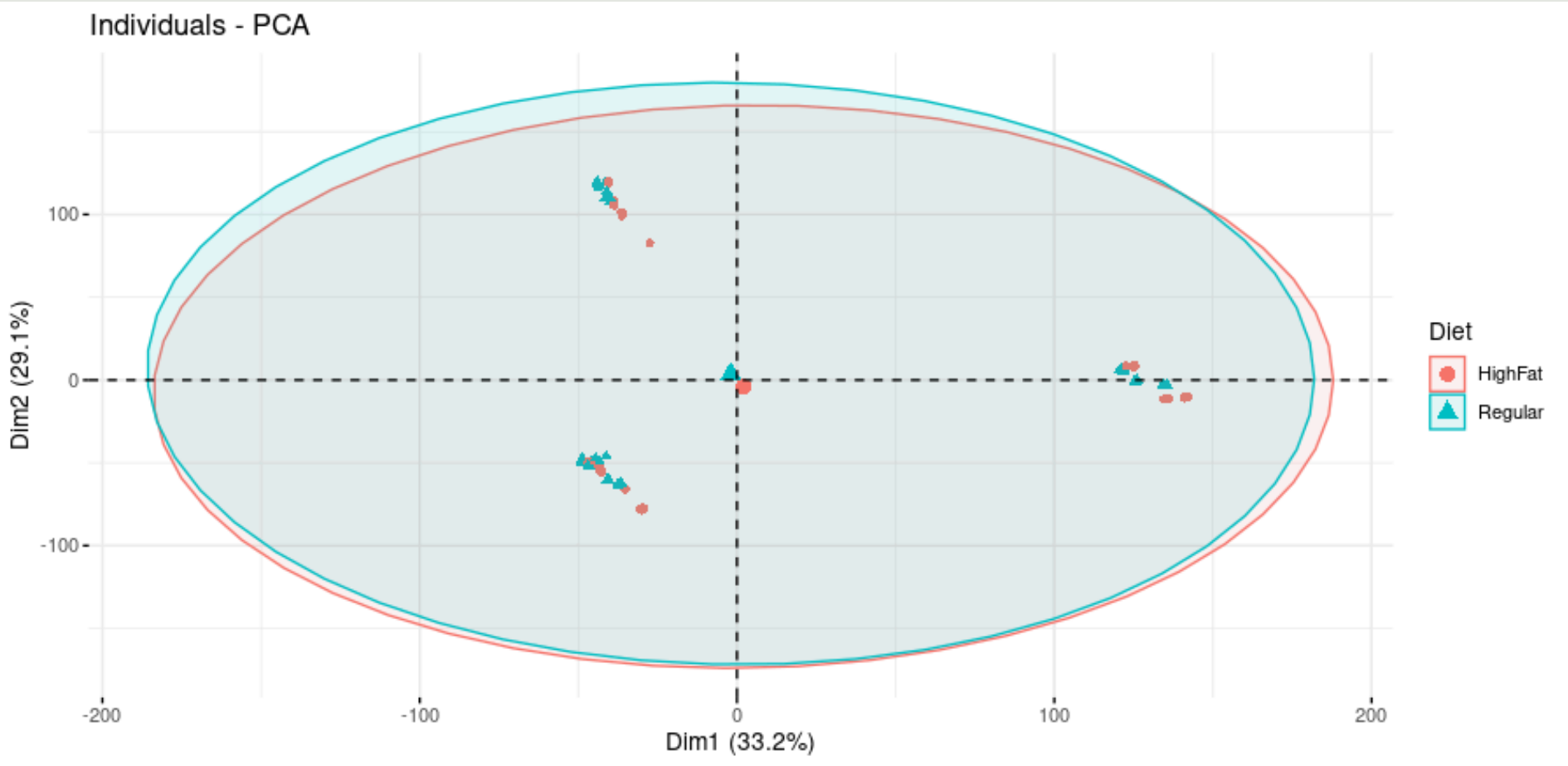
tSNE : BATCH EFFECTS

UMAP VISUALIZATION OF SAMPLE CLUSTERING BY SEX AND DIET
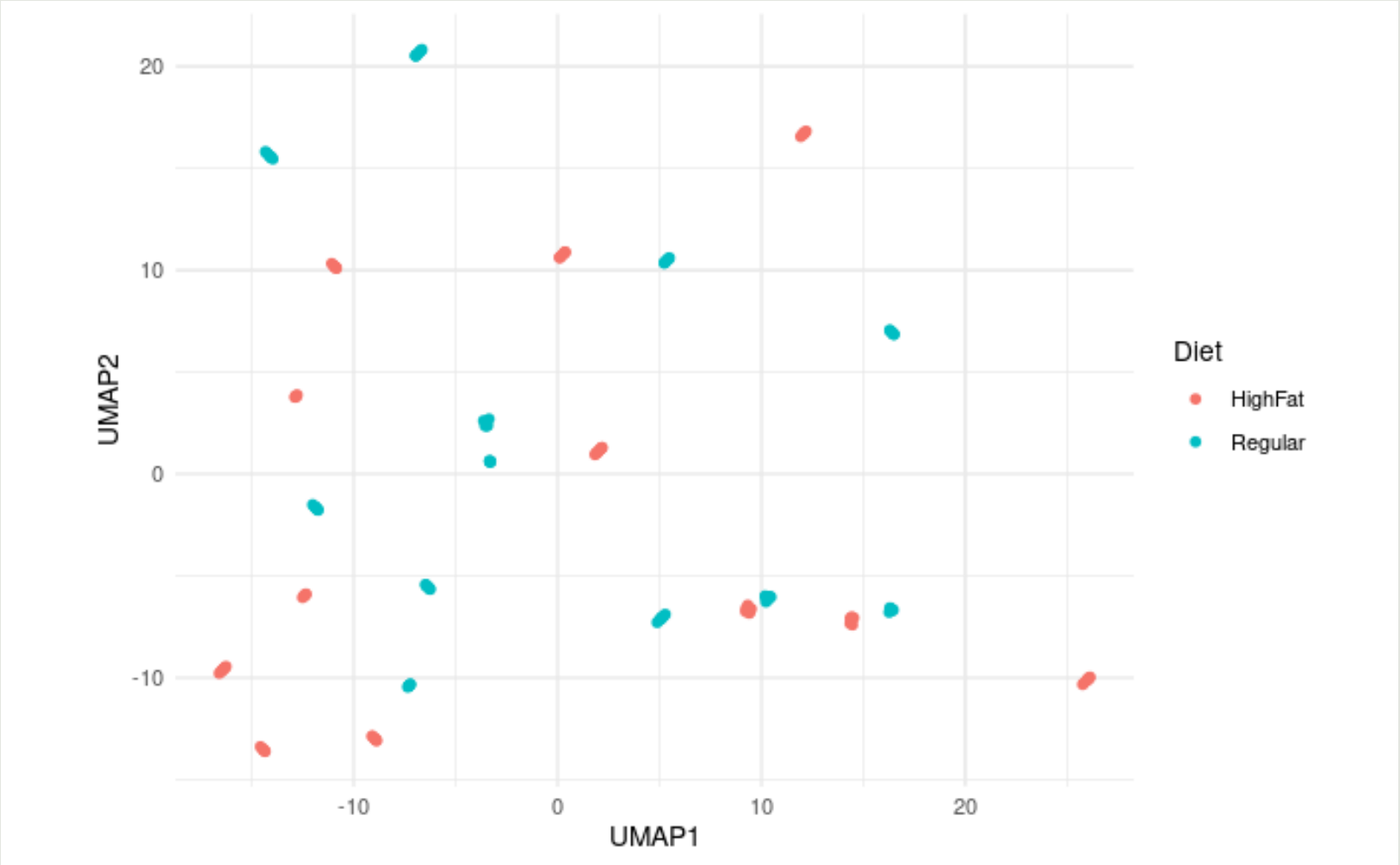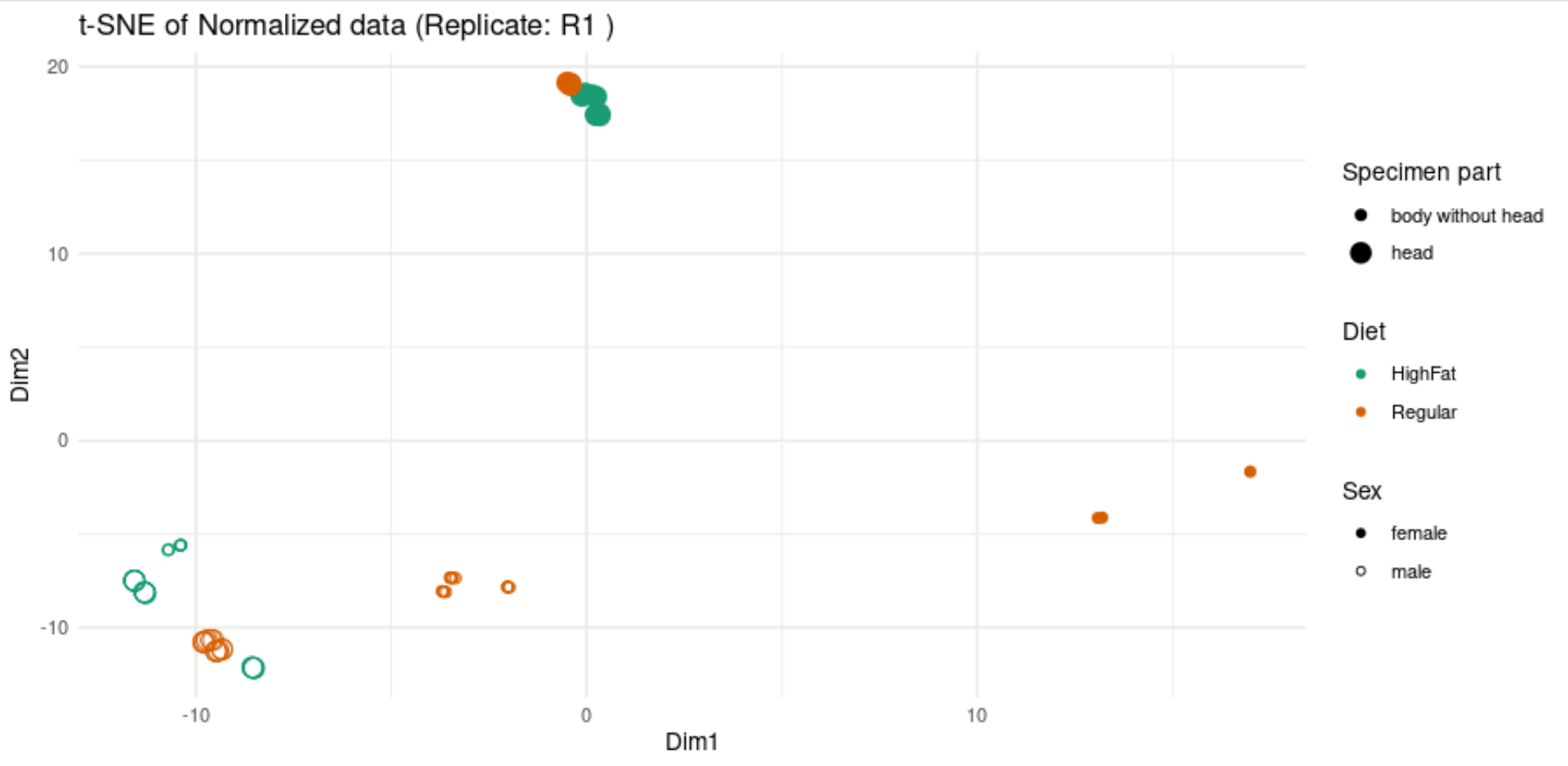

UMAP reveals distinct clustering based on sex, indicating that sex is a primary driver of variance in the dataset

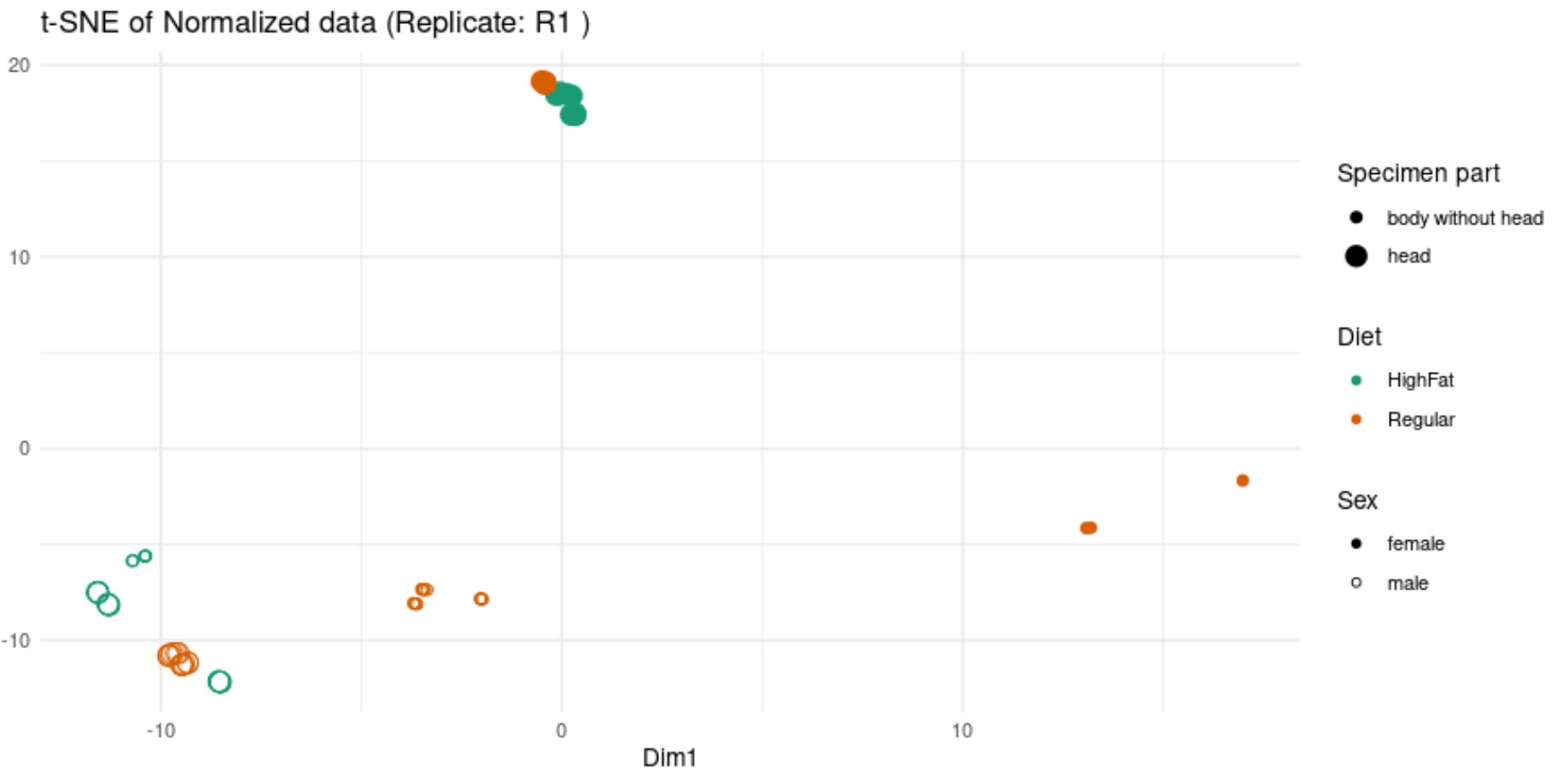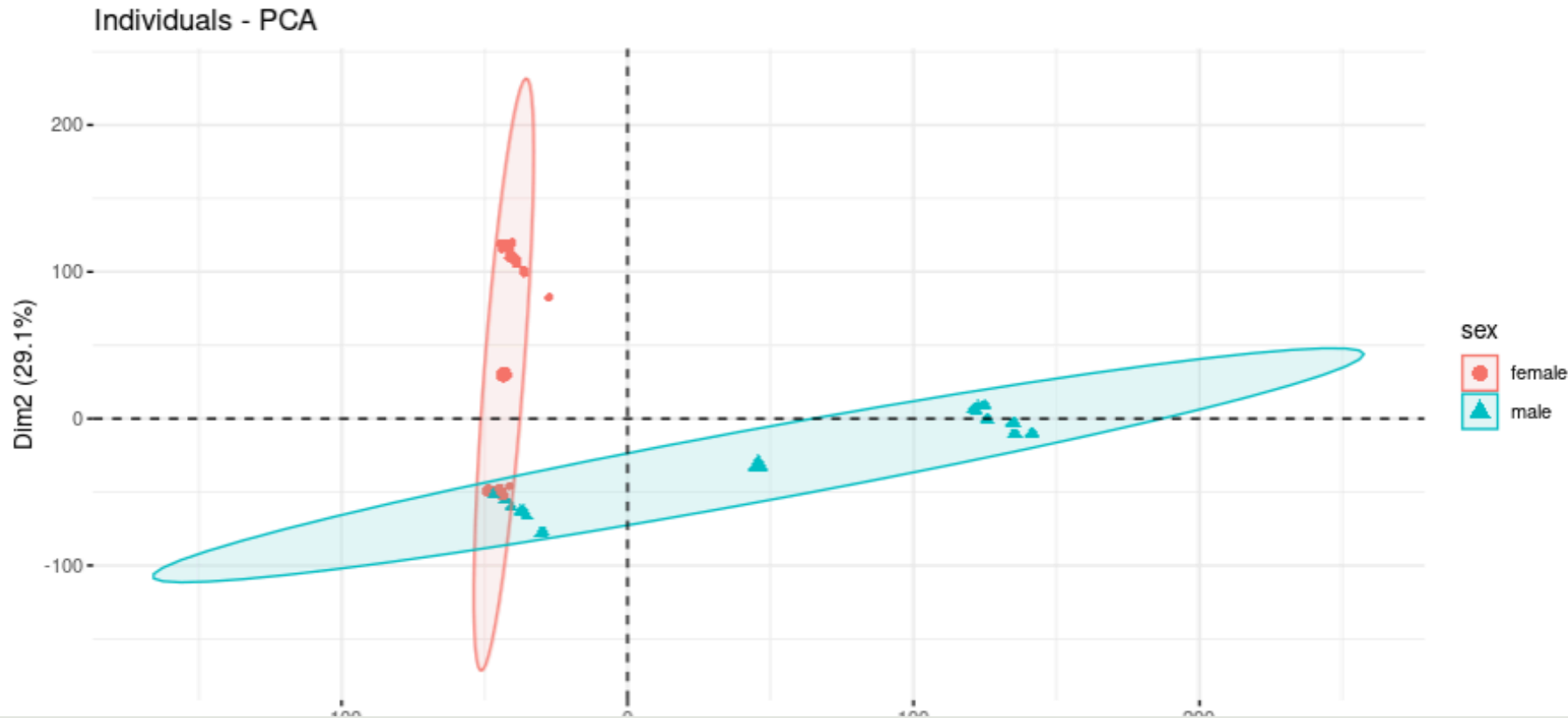No distinct clustering is observed by diet type, suggesting diet is not a major source of variance.
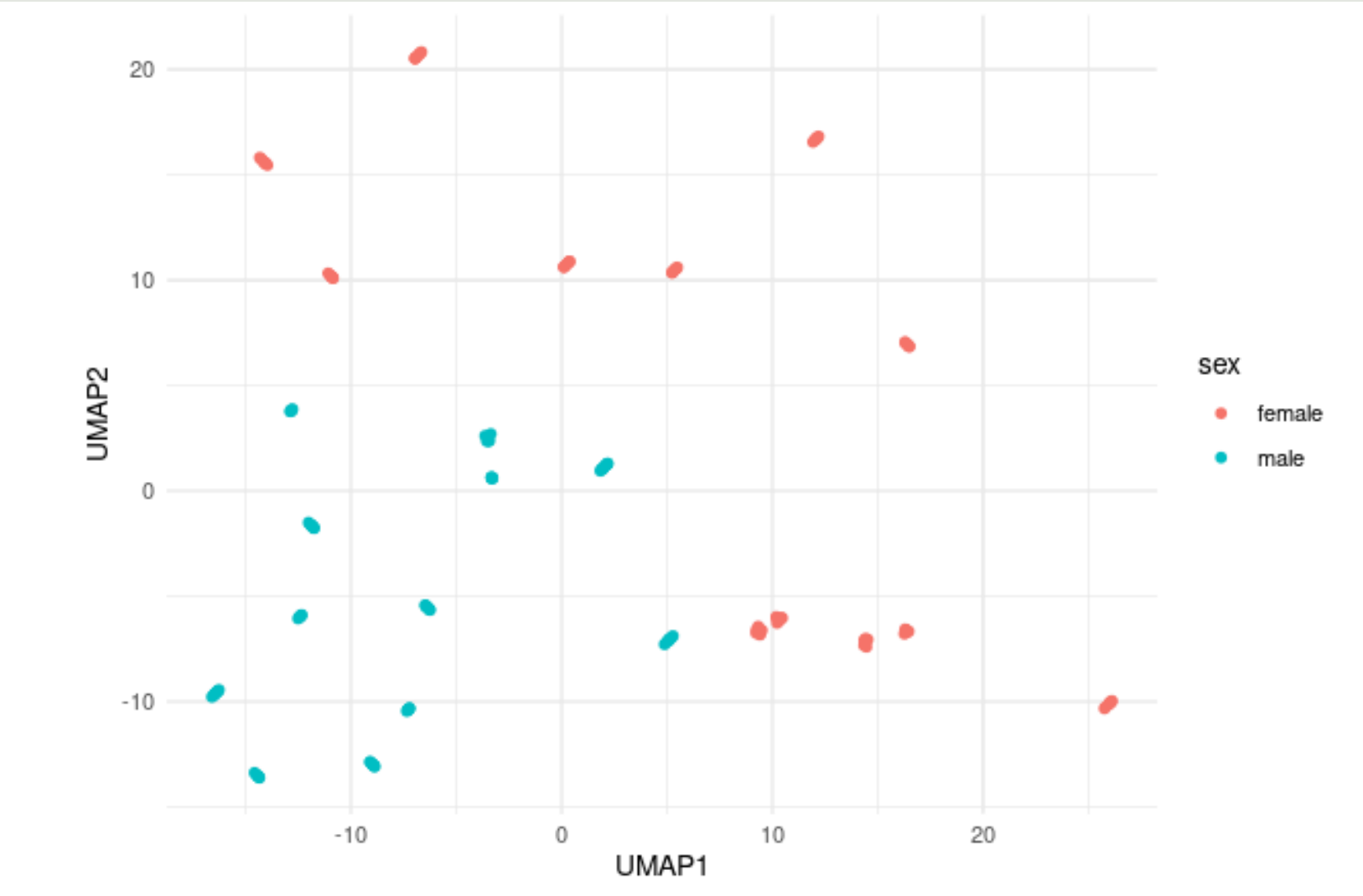
Conclusion of non-significant data

Conclusion of significant data

CONCLUSIONS

- Sex and body part drive gene expression variation.
- Sex-body part interaction explains PC1 variance.
- Diet (HFD vs. regular) has minor impact.
- Normalization removes biases, enhances biological patterns.
- No outliers or batch effects: data is consistent.
- PCA, t-SNE, UMAP show clusters by sex and body part..
- Reproducible results across parameter variations.
- Drosophila: valuable model for obesity and metabolic diseases.