**Linear models by individual KO sample, instead of merged, biggest coefficients highlighted:**

[1] "KO.G.1"

Call:

lm(formula = ns6Data[vars, i] ~ 0 + aveData[, 5:8])

Residuals:

Min 1Q Median 3Q Max

-3.7931 -0.2594 -0.0182 0.2533 2.9703

Coefficients:

Estimate Std. Error t value Pr(>|t|)

aveData[, 5:8]WT.G 0.970402 0.005578 173.985 < 2e-16 \*\*\*

aveData[, 5:8]WT.8 0.021356 0.011202 1.906 0.0567 .

aveData[, 5:8]WT.16 -0.170416 0.022502 -7.573 4.31e-14 \*\*\*

aveData[, 5:8]WT.24 0.175089 0.016504 10.609 < 2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.4757 on 4998 degrees of freedom

Multiple R-squared: 0.9977, Adjusted R-squared: 0.9977

F-statistic: 5.383e+05 on 4 and 4998 DF, p-value: < 2.2e-16

[1] "KO.G.2"

Call:

lm(formula = ns6Data[vars, i] ~ 0 + aveData[, 5:8])

Residuals:

Min 1Q Median 3Q Max

-3.8201 -0.3154 -0.0088 0.3223 3.8102

Coefficients:

Estimate Std. Error t value Pr(>|t|)

aveData[, 5:8]WT.G 0.995558 0.006692 148.766 < 2e-16 \*\*\*

aveData[, 5:8]WT.8 -0.043075 0.013441 -3.205 0.00136 \*\*

aveData[, 5:8]WT.16 -0.043624 0.026999 -1.616 0.10621

aveData[, 5:8]WT.24 0.088452 0.019803 4.467 8.12e-06 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.5708 on 4998 degrees of freedom

Multiple R-squared: 0.9967, Adjusted R-squared: 0.9967

F-statistic: 3.748e+05 on 4 and 4998 DF, p-value: < 2.2e-16

[1] "KO.G.3"

Call:

lm(formula = ns6Data[vars, i] ~ 0 + aveData[, 5:8])

Residuals:

Min 1Q Median 3Q Max

-4.0346 -0.2848 -0.0171 0.2791 2.6783

Coefficients:

Estimate Std. Error t value Pr(>|t|)

aveData[, 5:8]WT.G 0.923401 0.005787 159.551 <2e-16 \*\*\*

aveData[, 5:8]WT.8 0.106477 0.011624 9.160 <2e-16 \*\*\*

aveData[, 5:8]WT.16 -0.227768 0.023350 -9.755 <2e-16 \*\*\*

aveData[, 5:8]WT.24 0.192958 0.017126 11.267 <2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.4937 on 4998 degrees of freedom

Multiple R-squared: 0.9975, Adjusted R-squared: 0.9975

F-statistic: 4.986e+05 on 4 and 4998 DF, p-value: < 2.2e-16

[1] "KO.8.2"

Call:

lm(formula = ns6Data[vars, i] ~ 0 + aveData[, 5:8])

Residuals:

Min 1Q Median 3Q Max

-3.5538 -0.2935 0.0270 0.3287 3.3771

Coefficients:

Estimate Std. Error t value Pr(>|t|)

aveData[, 5:8]WT.G 0.152797 0.006134 24.91 <2e-16 \*\*\*

aveData[, 5:8]WT.8 0.863008 0.012319 70.05 <2e-16 \*\*\*

aveData[, 5:8]WT.16 -0.275702 0.024747 -11.14 <2e-16 \*\*\*

aveData[, 5:8]WT.24 0.254157 0.018151 14.00 <2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.5232 on 4998 degrees of freedom

Multiple R-squared: 0.9972, Adjusted R-squared: 0.9972

F-statistic: 4.435e+05 on 4 and 4998 DF, p-value: < 2.2e-16

[1] "KO.8.3"

Call:

lm(formula = ns6Data[vars, i] ~ 0 + aveData[, 5:8])

Residuals:

Min 1Q Median 3Q Max

-2.9434 -0.3377 0.0758 0.4370 3.4046

Coefficients:

Estimate Std. Error t value Pr(>|t|)

aveData[, 5:8]WT.G 0.471792 0.007997 59.00 <2e-16 \*\*\*

aveData[, 5:8]WT.8 0.983386 0.016061 61.23 <2e-16 \*\*\*

aveData[, 5:8]WT.16 -0.868375 0.032263 -26.91 <2e-16 \*\*\*

aveData[, 5:8]WT.24 0.406884 0.023664 17.20 <2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.6821 on 4998 degrees of freedom

Multiple R-squared: 0.9952, Adjusted R-squared: 0.9952

F-statistic: 2.616e+05 on 4 and 4998 DF, p-value: < 2.2e-16

[1] "KO.8.5"

Call:

lm(formula = ns6Data[vars, i] ~ 0 + aveData[, 5:8])

Residuals:

Min 1Q Median 3Q Max

-3.7983 -0.3152 0.0198 0.3445 3.4354

Coefficients:

Estimate Std. Error t value Pr(>|t|)

aveData[, 5:8]WT.G 0.457395 0.006535 69.988 <2e-16 \*\*\*

aveData[, 5:8]WT.8 0.413054 0.013126 31.469 <2e-16 \*\*\*

aveData[, 5:8]WT.16 0.095387 0.026367 3.618 0.0003 \*\*\*

aveData[, 5:8]WT.24 0.031656 0.019339 1.637 0.1017

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.5574 on 4998 degrees of freedom

Multiple R-squared: 0.9968, Adjusted R-squared: 0.9968

F-statistic: 3.921e+05 on 4 and 4998 DF, p-value: < 2.2e-16

[1] "KO.16.1"

Call:

lm(formula = ns6Data[vars, i] ~ 0 + aveData[, 5:8])

Residuals:

Min 1Q Median 3Q Max

-3.6425 -0.4152 -0.0324 0.3858 3.2951

Coefficients:

Estimate Std. Error t value Pr(>|t|)

aveData[, 5:8]WT.G 0.618132 0.007448 82.993 <2e-16 \*\*\*

aveData[, 5:8]WT.8 0.184084 0.014959 12.306 <2e-16 \*\*\*

aveData[, 5:8]WT.16 0.020628 0.030049 0.686 0.492

aveData[, 5:8]WT.24 0.184370 0.022039 8.365 <2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.6353 on 4998 degrees of freedom

Multiple R-squared: 0.9959, Adjusted R-squared: 0.9959

F-statistic: 3.068e+05 on 4 and 4998 DF, p-value: < 2.2e-16

[1] "KO.16.2"

Call:

lm(formula = ns6Data[vars, i] ~ 0 + aveData[, 5:8])

Residuals:

Min 1Q Median 3Q Max

-4.0540 -0.4655 0.0397 0.5400 3.8506

Coefficients:

Estimate Std. Error t value Pr(>|t|)

aveData[, 5:8]WT.G 0.41737 0.00951 43.88 <2e-16 \*\*\*

aveData[, 5:8]WT.8 0.78775 0.01910 41.24 <2e-16 \*\*\*

aveData[, 5:8]WT.16 -0.88929 0.03837 -23.18 <2e-16 \*\*\*

aveData[, 5:8]WT.24 0.68068 0.02814 24.19 <2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.8112 on 4998 degrees of freedom

Multiple R-squared: 0.9933, Adjusted R-squared: 0.9933

F-statistic: 1.846e+05 on 4 and 4998 DF, p-value: < 2.2e-16

[1] "KO.16.3"

Call:

lm(formula = ns6Data[vars, i] ~ 0 + aveData[, 5:8])

Residuals:

Min 1Q Median 3Q Max

-4.4804 -0.2686 0.0140 0.2935 2.9995

Coefficients:

Estimate Std. Error t value Pr(>|t|)

aveData[, 5:8]WT.G 0.20604 0.00550 37.459 < 2e-16 \*\*\*

aveData[, 5:8]WT.8 0.22281 0.01105 20.169 < 2e-16 \*\*\*

aveData[, 5:8]WT.16 0.63104 0.02219 28.436 < 2e-16 \*\*\*

aveData[, 5:8]WT.24 -0.06376 0.01628 -3.917 9.07e-05 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.4692 on 4998 degrees of freedom

Multiple R-squared: 0.9977, Adjusted R-squared: 0.9977

F-statistic: 5.51e+05 on 4 and 4998 DF, p-value: < 2.2e-16

[1] "KO.24.1"

Call:

lm(formula = ns6Data[vars, i] ~ 0 + aveData[, 5:8])

Residuals:

Min 1Q Median 3Q Max

-3.5364 -0.2472 0.0791 0.3480 2.6411

Coefficients:

Estimate Std. Error t value Pr(>|t|)

aveData[, 5:8]WT.G 0.122490 0.005835 20.991 <2e-16 \*\*\*

aveData[, 5:8]WT.8 -0.026030 0.011720 -2.221 0.0264 \*

aveData[, 5:8]WT.16 0.441832 0.023543 18.767 <2e-16 \*\*\*

aveData[, 5:8]WT.24 0.452119 0.017267 26.184 <2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.4977 on 4998 degrees of freedom

Multiple R-squared: 0.9974, Adjusted R-squared: 0.9974

F-statistic: 4.821e+05 on 4 and 4998 DF, p-value: < 2.2e-16

[1] "KO.24.2"

Call:

lm(formula = ns6Data[vars, i] ~ 0 + aveData[, 5:8])

Residuals:

Min 1Q Median 3Q Max

-3.2087 -0.3617 0.0362 0.4086 2.9529

Coefficients:

Estimate Std. Error t value Pr(>|t|)

aveData[, 5:8]WT.G 0.094314 0.006986 13.500 < 2e-16 \*\*\*

aveData[, 5:8]WT.8 0.295163 0.014031 21.036 < 2e-16 \*\*\*

aveData[, 5:8]WT.16 0.120877 0.028186 4.289 1.83e-05 \*\*\*

aveData[, 5:8]WT.24 0.488278 0.020673 23.619 < 2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.5959 on 4998 degrees of freedom

Multiple R-squared: 0.9964, Adjusted R-squared: 0.9964

F-statistic: 3.424e+05 on 4 and 4998 DF, p-value: < 2.2e-16

[1] "KO.24.3"

Call:

lm(formula = ns6Data[vars, i] ~ 0 + aveData[, 5:8])

Residuals:

Min 1Q Median 3Q Max

-3.7690 -0.4174 -0.0223 0.3765 3.1583

Coefficients:

Estimate Std. Error t value Pr(>|t|)

aveData[, 5:8]WT.G 0.388139 0.007302 53.15 <2e-16 \*\*\*

aveData[, 5:8]WT.8 0.607554 0.014666 41.42 <2e-16 \*\*\*

aveData[, 5:8]WT.16 -0.435778 0.029461 -14.79 <2e-16 \*\*\*

aveData[, 5:8]WT.24 0.443856 0.021608 20.54 <2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.6229 on 4998 degrees of freedom

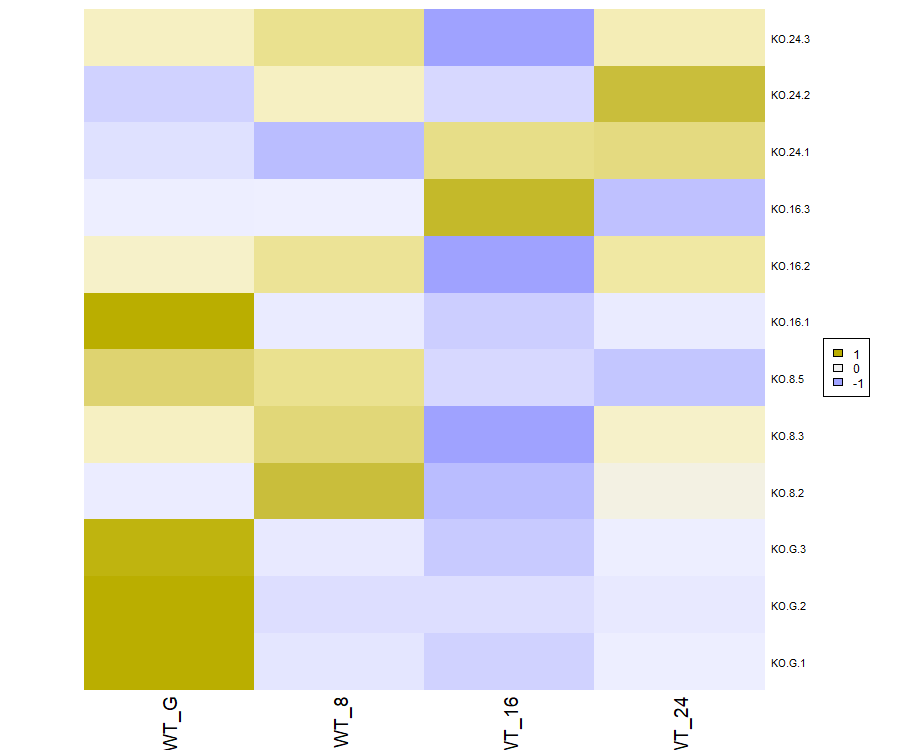
Multiple R-squared: 0.9961, Adjusted R-squared: 0.9961

F-statistic: 3.173e+05 on 4 and 4998 DF, p-value: < 2.2e-16

**The coefficients from the above linear models:**

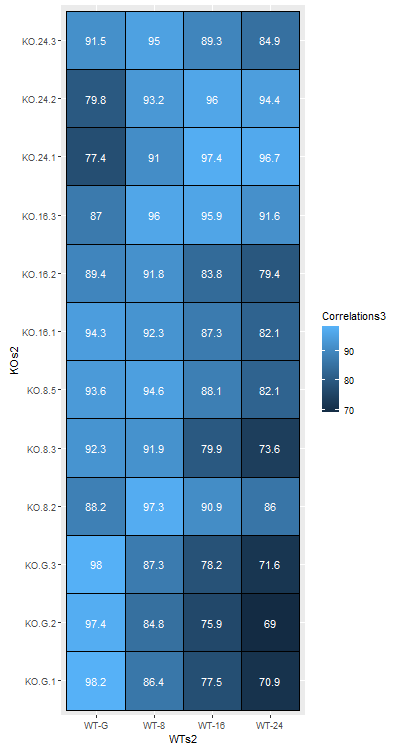
This is a little weird to ‘read’ in places since e.g. the KO\_24\_3 sample has its highest expression in the negative, corresponding to WT\_16. Maybe I should flip the sign on any where the highest is a negative, or give it the same treatment as the correlation stuff below.

All three KO\_G are described best by WT\_G. One of each group KO\_8 and KO\_16 are falling behind, appearing most similar to WT\_G. One KO\_24 sample was most similar to WT\_8.



**Pair-wise Correlations for each KO sample:**

More consistently showing samples arrested in development.



**PCA and glmpca()**

GLMPCA showing the same story as PCA really, although I didn’t factor in the batch effect with glmpca since we aren’t sure yet that there really is one experimentally speaking – I think Scott is getting back to us on that. Development goes with PC1 in PCA and with dim2 (maybe dim1 as well?) in glmPCA. Dim1 in particular seems to separate the gametocytes from ookinetes, where dim2 has better separation of ookinete time points. Should I just stick to PCA since that is what I already started talking to Scott about or should I share/use this too?

