Exploration of Multiple Treatments on the Metabolic Efficiency of the Mitochondria

Dom Fenoglio, Brian Kim, Kathleen Zhang

1 Background

The mitochondria are considered the "powerhouse" of the cell, responsible for generating the cell's usable energy through oxidative phosphorylation, a process necessary for all biological processes, particularly in high-demand organs such as the heart, brain, and muscles. Their proper function is critical for overall health, as disruptions to their function are associated with various health issues, such as cancer, heart disease, and Alzheimer's.

One way to examine mitochondrial function is by using the multiplexed assay platform, a laboratory method that allows researchers to measure multiple dimensions of mitochondrial activity across different substrates and energy demand conditions. By measuring respiration rates under different combinations of substrates across different experimental settings, such as genetic background and dose, researchers hope to better understand these effects on the metabolic and functional phenotypes of mitochondria.

The main motivation for our analysis is to quantitatively test hypotheses about genetic changes on mitochondrial efficiency and energy production, and whether there is evidence that genotype effects (transgenic vs. natural mice) depend on substrate and/or dose. By building a modeling framework, we hope to determine how mitochondrial efficiency varies by substrate, genotype, and dose while capturing both fixed and random sources of variation.

2 Exploratory Data Analysis

We explored patterns of VO_2 production across genotype, substrate, and dose to assess whether systematic differences exist.

Looking at figure 1 we can see that genotype has an effect on VO_2 production. Across nearly all substrates we can see the transgenic mice display higher VO_2 production than natural mice. For substrates of $PMOc$ and $PMPc$ we can see different slopes for VO_2 production vs Dose and as the doses become higher, the effects of genotype becomes more significant. This suggests a need for an interaction between dose and substrate. The effect is most pronounced

when the doses are higher in $PMPc$ and $PMOc$, and for OcM and PcM we see a clear higher VO_2 production for all doses. We also can see that substrates involving Octanoyl Carnitine Oc and Palmitoyl-Carnitine Pc has a more pronounced separation between transgenic vs natural mouse.

OcM and PcM show a relatively flat dose-response curves for both genotypes which suggests limited sensitivity to dose changes and $PMOc$ and $PMPc$ substrates highlight a stronger genotype effect as the transgenic mice has a more pronounced effect to dose. This points to an interaction between substrate and genotype where certain substrates amplify the genotype-specific differences in the VO_2 production efficiency.

Pair - Level Variation

The researchers' experimental design, which matched a transgenic mouse with a natural type mouse and tested each pair on a different day, could induce some added variation that dose and substrate cannot account for. This is because the experimental setup could vary slightly day-to-day, influencing the measurement of our response variable VO_2.

Looking at Figure 2, we can see that there does in fact seem to be systematic differences on the pair level. For example, pair 5 exhibits a much larger gap between VO_2 production of transgenic and natural type mice for OcM and PcM. Across every pair, we see enough variation between the genotype-VO_2 relationship to warrant consideration in our final modeling decisions.

Conclusions

1. Genotype appears significantly correlated with VO_2 production: transgenic mice had higher VO_2 than nautral mice across most conditions
2. Substrate and dose seem to matter as well: We saw certain substrates had a larger effect in the VO_2 production as $PMOc$ and $PMPc$ tended to effect the transgenic mice more while OcM and PcM tended to have an equal effect on transgenic and natural mice. This relationship was highly dependent on dose.
3. Pairs: While researchers do not need to know the specific measurement predictions for each pair, we need to account for variation on the pair level in our model.

3 Modeling

To account for all of the points noted above, we chose to fit a fully interactive linear model regressing genotype, dose, and substrate on VO_2. Furthermore, we included a random intercept for pair, allowing us to include this added variation in the model while retaining the ability to predict on an unobserved pair.

Comparing this model to two other models – a model excluding genotype and a model excluding the random intercept – we see that our chosen equation is the best fit for the data.

$$Model\ 1:\ VO_2\ natural * Dose * Substrate + (1|pair)$$
$$Model\ 2:\ VO_2\ natural * Dose * Substrate$$
$$Model\ 3:\ VO_2\ Dose * Substrate + (1|pair)$$

From Table 1, we can see a log likelihood value of -2489.7 for our chosen model, and lower values for model 2 and 3 (-2759.2 and -2724.9, respectively). This supports the conclusion that our chosen model captures the behavior of the data the best.

NEED TO FIX p-values? Looking at the chi squared and the p value which compare the model fit of model 2 vs 1 and model 3 vs 1 we can see that the p values are very small (close to 0). Since the p values are very small the improvement in model 1 is not by chance so in model 2 it shows the random intercept for pair is essential. The p value comparing model 3 and 1 is also very small (close to 0) which suggests that genotype is also significant. Since the p values are very small the worse fit of model 2 and 3 is statistically significant which indicates their ommitted terms are necessary.

Furthermore, Table 2 shows our models intraclass correlation (ICC), which is the ratio of pair variability to total variability. An ICC value of 0.644 is a strong piece of evidence in support of including a random intercept for pair, as it signifies that the variation across pairs is non-neglible.

The conditional R^2 value in Table 2 represents the variance explained by our model. By allowing each pair have its own intercept we explain about 96.5% of the total variability in our data, higher than the 90% explained marginally by the fixed effects.

4 Analysis

TODO: bonferroni correction for p-values, include/analyze graph and give interpretation examples (see presentation).

5 Conclusion and Future Work

Our analysis showed strong evidence that genotype significantly influences VO_2 production conditional on both substrate and dose. Across nearly all experimental conditions, transgenic mice displayed a higher VO_2 production relative to natural mice with some substrates amplifying this effect more than others. These findings show that transgenic geotype is associated with enhanced metabolic efficiency

However, one major limitation can be seen in Figure 4, the residual plot across dosage levels. We see some evidence of a nonlinear relationship between dose and VO_2 production conditional on substrate and genotype, something that was not accounted for in our model. We chose to live the relationship linear because it was our understanding that the researchers' had some biological motivation behind this claim. Since our analysis shows evidence arguing against this claim, further work should include interrogating these assumptions, especially among higher dosage levels.

In addition, Future work should focus on exploring a broader range of substrates and leveraging larger samples to better account for variability across experimental pairs. These extensions would help clarify the extent to which the observed genotype effects generalize across different biological and experimental contexts.
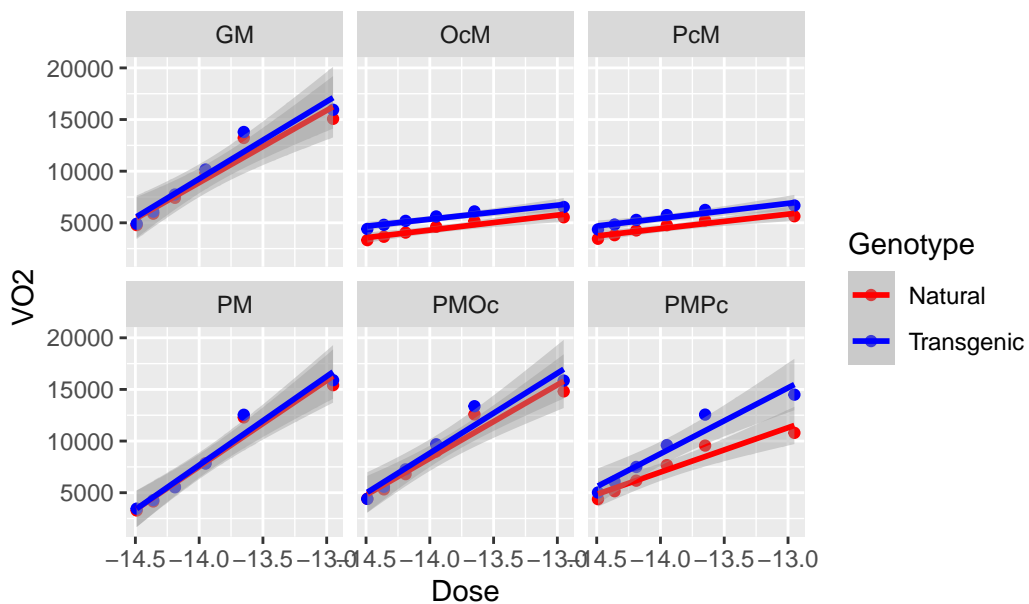
6 Appendix

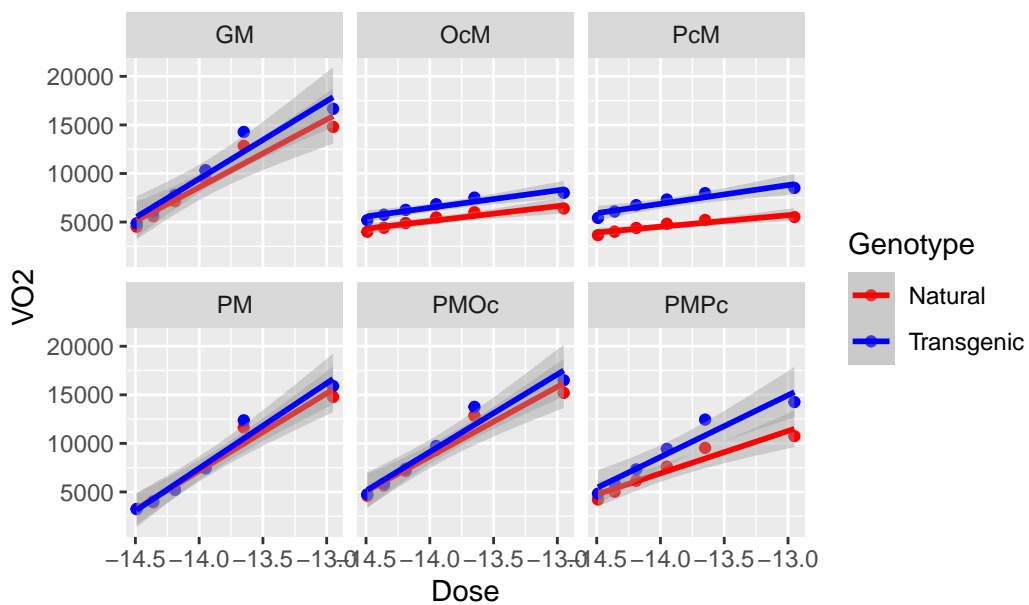**Exploratory Data Analysis**
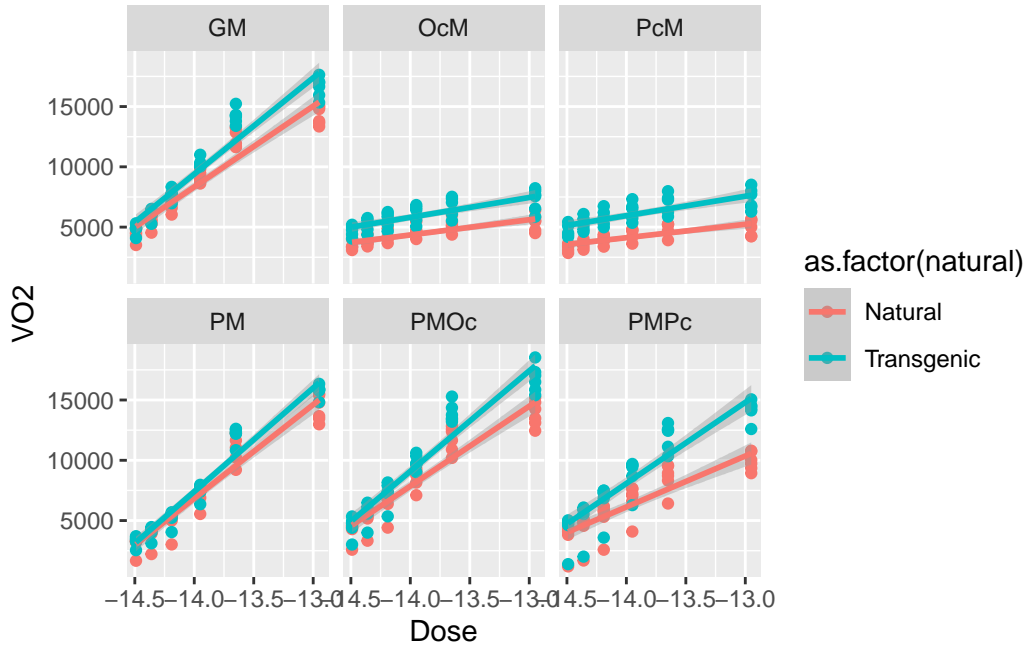
VO2 vs. Dose for pair 1

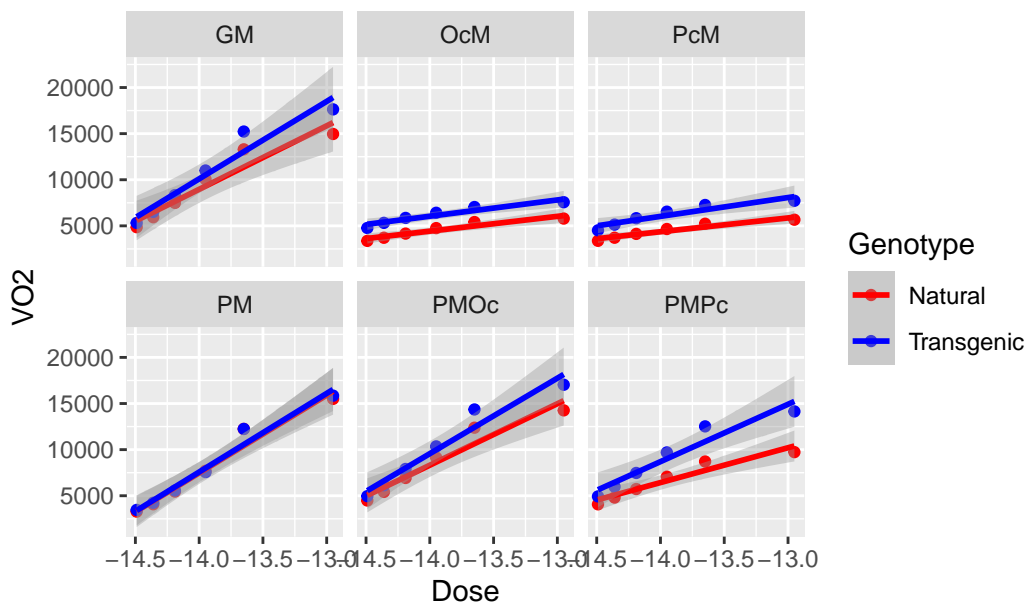VO2 vs. Dose for pair 2

VO2 vs. Dose for pair 3
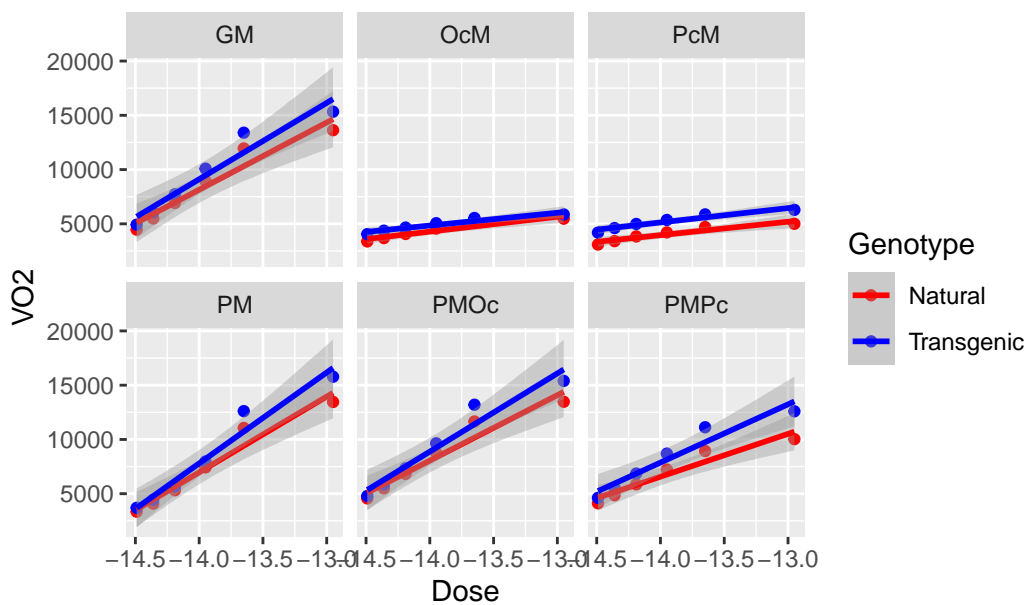
VO2 vs. Dose for pair 4

VO2 vs. Dose for pair 5
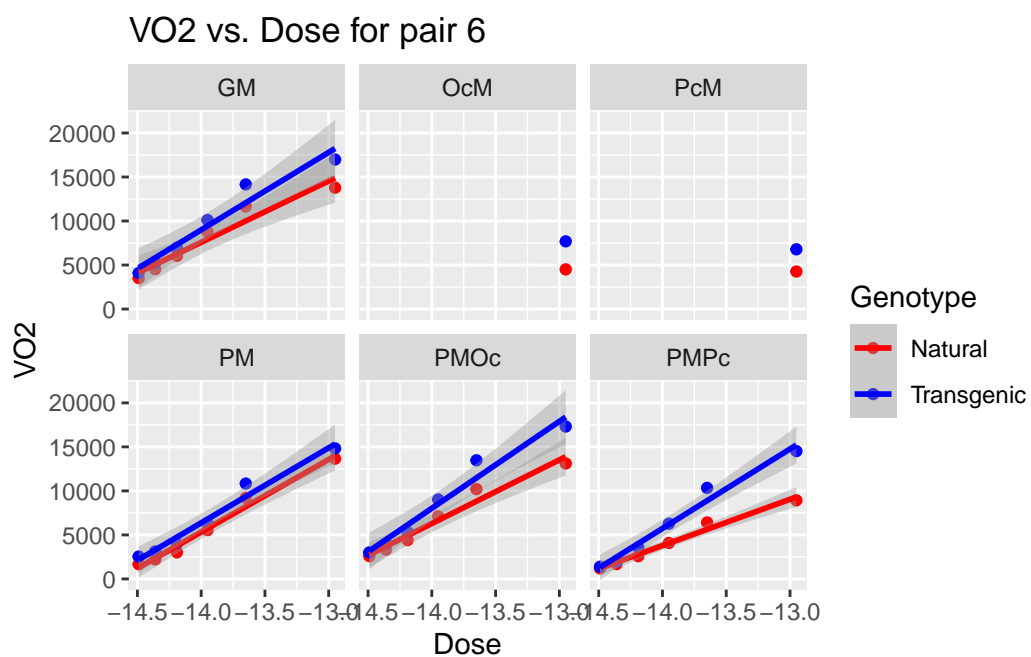

VO2 vs. Dose for pair 6
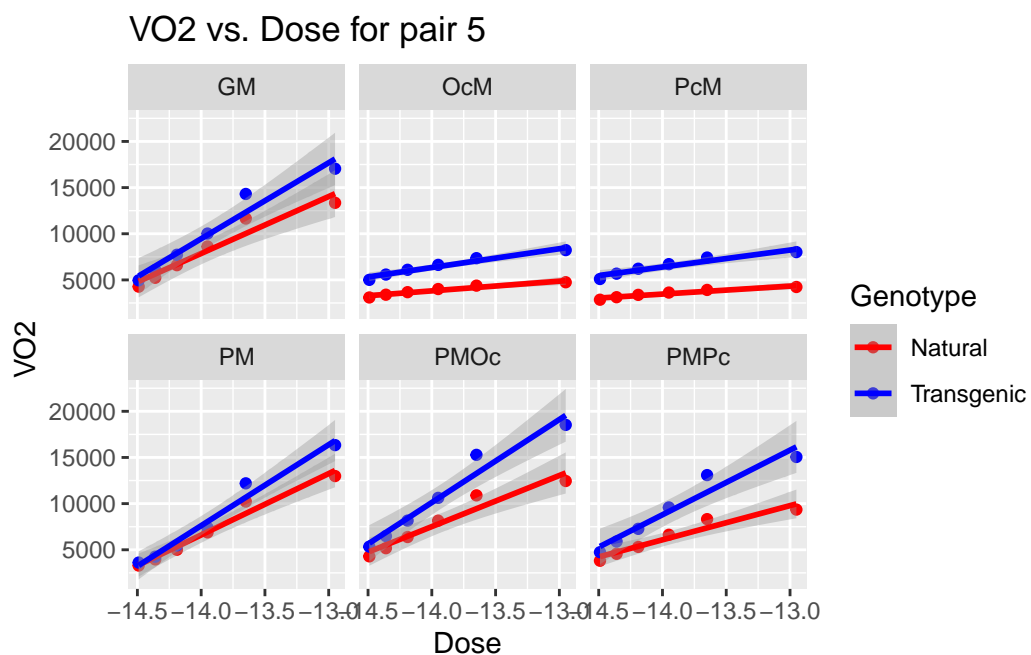
Table 1: Model comparison using AIC and Adjusted R²

Model	df	AIC	Adjusted_R2
Amino Acids	8	7439.921	0.721

Substrate	9	7386.290	0.755

Modeling

```
Call:
lm(formula = VO2 ~ Substrate:natural + Dose:natural:Substrate,
    data = data)

Residuals:
    Min      1Q  Median      3Q     Max
-3254.1  -370.8   186.2   536.4  1727.3

Coefficients: (1 not defined because of singularities)
                                   Estimate Std. Error t value Pr(>|t|)
(Intercept)                         138201.0     7230.7  19.113  < 2e-16 ***
SubstrateGM:naturalNatural            5077.4    10225.7   0.497 0.619863
SubstrateOcM:naturalNatural        -104126.4    10724.8  -9.709  < 2e-16 ***
SubstratePcM:naturalNatural        -107896.1    10724.8 -10.060  < 2e-16 ***
SubstratePM:naturalNatural            3735.0    10225.7   0.365 0.715163
SubstratePMOc:naturalNatural         -2949.2    10225.7  -0.288 0.773223
SubstratePMPc:naturalNatural        -49273.5    10225.7  -4.819 2.25e-06 ***
SubstrateGM:naturalTransgenic        27374.7    10225.7   2.677 0.007815 **
SubstrateOcM:naturalTransgenic      -99396.5    10724.8  -9.268  < 2e-16 ***
SubstratePcM:naturalTransgenic      -95812.6    10724.8  -8.934  < 2e-16 ***
SubstratePM:naturalTransgenic        14409.0    10225.7   1.409 0.159790
SubstratePMOc:naturalTransgenic      27257.2    10225.7   2.666 0.008081 **
SubstratePMPc:naturalTransgenic           NA         NA      NA       NA
SubstrateGM:naturalNatural:Dose       9596.5      511.7  18.755  < 2e-16 ***
SubstrateOcM:naturalNatural:Dose      2110.8      560.5   3.766 0.000198 ***
SubstratePcM:naturalNatural:Dose      1859.4      560.5   3.317 0.001015 **
SubstratePM:naturalNatural:Dose       9622.6      511.7  18.806  < 2e-16 ***
SubstratePMOc:naturalNatural:Dose     9064.0      511.7  17.714  < 2e-16 ***
SubstratePMPc:naturalNatural:Dose     5889.9      511.7  11.511  < 2e-16 ***
SubstrateGM:naturalTransgenic:Dose   11110.1      511.7  21.713  < 2e-16 ***
SubstrateOcM:naturalTransgenic:Dose   2346.2      560.5   4.186 3.69e-05 ***
SubstratePcM:naturalTransgenic:Dose   2588.6      560.5   4.618 5.64e-06 ***
SubstratePM:naturalTransgenic:Dose   10345.9      511.7  20.219  < 2e-16 ***
SubstratePMOc:naturalTransgenic:Dose 11128.3      511.7  21.748  < 2e-16 ***
SubstratePMPc:naturalTransgenic:Dose  9257.5      511.7  18.092  < 2e-16 ***
---
```

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 839.7 on 316 degrees of freedom
  (20 observations deleted due to missingness)
Multiple R-squared: 0.9233,   Adjusted R-squared: 0.9177
F-statistic: 165.4 on 23 and 316 DF, p-value: < 2.2e-16

Call:
lm(formula = VO2 ~ natural * Dose * Substrate, data = data)

Residuals:
    Min   1Q Median   3Q   Max
-3254.1 -370.8   186.2   536.4 1727.3

Coefficients:
	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	143278.42	7230.68	19.815	< 2e-16	***
naturalTransgenic	22297.32	10225.72	2.181	0.0300	*
Dose	9596.47	511.68	18.755	< 2e-16	***
SubstrateOcM	-109203.79	10724.83	-10.182	< 2e-16	***
SubstratePcM	-112973.48	10724.83	-10.534	< 2e-16	***
SubstratePM	-1342.39	10225.72	-0.131	0.8956	
SubstratePMOc	-8026.61	10225.72	-0.785	0.4331	
SubstratePMPc	-54350.93	10225.72	-5.315	2.02e-07	***
naturalTransgenic:Dose	1513.62	723.63	2.092	0.0373	*
naturalTransgenic:SubstrateOcM	-17567.40	15167.20	-1.158	0.2476	
naturalTransgenic:SubstratePcM	-10213.84	15167.20	-0.673	0.5012	
naturalTransgenic:SubstratePM	-11623.31	14461.35	-0.804	0.4221	
naturalTransgenic:SubstratePMOc	7909.05	14461.35	0.547	0.5848	
naturalTransgenic:SubstratePMPc	26976.20	14461.35	1.865	0.0631	.
Dose:SubstrateOcM	-7485.71	758.95	-9.863	< 2e-16	***
Dose:SubstratePcM	-7737.11	758.95	-10.195	< 2e-16	***
Dose:SubstratePM	26.09	723.63	0.036	0.9713	
Dose:SubstratePMOc	-532.46	723.63	-0.736	0.4624	
Dose:SubstratePMPc	-3706.61	723.63	-5.122	5.26e-07	***
naturalTransgenic:Dose:SubstrateOcM	-1278.14	1073.31	-1.191	0.2346	
naturalTransgenic:Dose:SubstratePcM	-784.42	1073.31	-0.731	0.4654	
naturalTransgenic:Dose:SubstratePM	-790.30	1023.37	-0.772	0.4405	
naturalTransgenic:Dose:SubstratePMOc	550.68	1023.37	0.538	0.5909	
naturalTransgenic:Dose:SubstratePMPc	1854.02	1023.37	1.812	0.0710	.

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 839.7 on 316 degrees of freedom
  (20 observations deleted due to missingness)
Multiple R-squared: 0.9233,   Adjusted R-squared: 0.9177
F-statistic: 165.4 on 23 and 316 DF, p-value: < 2.2e-16

Call:
lm(formula = VO2 ~ Substrate - 1, data = data)

Residuals:
    Min   1Q Median   3Q   Max
-5366.1 -1939.5 -445.7 1327.1 7567.6

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
SubstrateGM   8155.9   346.4   23.55   <2e-16 ***
SubstrateOcM   4955.3   379.4   13.06   <2e-16 ***
SubstratePcM   4926.5   379.4   12.98   <2e-16 ***
SubstratePM   6215.9   346.4   17.95   <2e-16 ***
SubstratePMOc 7716.4   346.4   22.28   <2e-16 ***
SubstratePMPc 6563.2   346.4   18.95   <2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2683 on 334 degrees of freedom
  (20 observations deleted due to missingness)
Multiple R-squared: 0.8611,   Adjusted R-squared: 0.8586
F-statistic: 345.1 on 6 and 334 DF, p-value: < 2.2e-16

Linear mixed model fit by REML. t-tests use Satterthwaite's method [
lmerModLmerTest]
Formula: VO2 ~ natural * Dose * Substrate + (1 | pair)
   Data: data

REML criterion at convergence: 4979.4

Scaled residuals:
     Min   1Q   Median   3Q   Max
-3.09075 -0.64782 0.02403 0.63668 3.05598

```
Random effects:
 Groups    Name         Variance Std.Dev.
 pair      (Intercept)  582983   763.5
 Residual               322707   568.1
Number of obs: 340, groups:  pair, 6

Fixed effects:
                                           Estimate Std. Error        df t value
(Intercept)                                143278.42    4901.42    313.21  29.232
naturalTransgenic                           22297.32    6917.62    310.97   3.223
Dose                                         9596.47     346.15    310.97  27.724
SubstrateOcM                              -109503.46    7255.28    310.97 -15.093
SubstratePcM                              -113273.15    7255.28    310.97 -15.613
SubstratePM                                 -1342.39    6917.62    310.97  -0.194
SubstratePMOc                               -8026.61    6917.62    310.97  -1.160
SubstratePMPc                              -54350.93    6917.62    310.97  -7.857
naturalTransgenic:Dose                       1513.62     489.53    310.97   3.092
naturalTransgenic:SubstrateOcM             -17567.40   10260.49    310.97  -1.712
naturalTransgenic:SubstratePcM             -10213.84   10260.49    310.97  -0.995
naturalTransgenic:SubstratePM              -11623.31    9782.99    310.97  -1.188
naturalTransgenic:SubstratePMOc              7909.05    9782.99    310.97   0.808
naturalTransgenic:SubstratePMPc             26976.20    9782.99    310.97   2.757
Dose:SubstrateOcM                           -7485.71     513.42    310.97 -14.580
Dose:SubstratePcM                           -7737.11     513.42    310.97 -15.070
Dose:SubstratePM                               26.09     489.53    310.97   0.053
Dose:SubstratePMOc                           -532.46     489.53    310.97  -1.088
Dose:SubstratePMPc                          -3706.61     489.53    310.97  -7.572
naturalTransgenic:Dose:SubstrateOcM         -1278.14     726.09    310.97  -1.760
naturalTransgenic:Dose:SubstratePcM          -784.42     726.09    310.97  -1.080
naturalTransgenic:Dose:SubstratePM           -790.30     692.30    310.97  -1.142
naturalTransgenic:Dose:SubstratePMOc          550.68     692.30    310.97   0.795
naturalTransgenic:Dose:SubstratePMPc         1854.02     692.30    310.97   2.678
                                           Pr(>|t|)
(Intercept)                                < 2e-16 ***
naturalTransgenic                          0.00140 **
Dose                                       < 2e-16 ***
SubstrateOcM                               < 2e-16 ***
SubstratePcM                               < 2e-16 ***
SubstratePM                                0.84626
SubstratePMOc                              0.24681
SubstratePMPc                              6.48e-14 ***
naturalTransgenic:Dose                     0.00217 **
naturalTransgenic:SubstrateOcM             0.08787 .
```

```
naturalTransgenic:SubstratePcM         0.32029
naturalTransgenic:SubstratePM          0.23570
naturalTransgenic:SubstratePMOc        0.41945
naturalTransgenic:SubstratePMPc        0.00617 **
Dose:SubstrateOcM                     < 2e-16 ***
Dose:SubstratePcM                     < 2e-16 ***
Dose:SubstratePM                       0.95753
Dose:SubstratePMOc                     0.27757
Dose:SubstratePMPc                    4.23e-13 ***
naturalTransgenic:Dose:SubstrateOcM    0.07934 .
naturalTransgenic:Dose:SubstratePcM    0.28083
naturalTransgenic:Dose:SubstratePM     0.25451
naturalTransgenic:Dose:SubstratePMOc   0.42697
naturalTransgenic:Dose:SubstratePMPc   0.00780 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Correlation matrix not shown by default, as p = 24 > 12.
Use print(x, correlation=TRUE)  or
    vcov(x)         if you need it


Type III Analysis of Variance Table with Satterthwaite's method
                         Sum Sq     Mean Sq NumDF  DenDF    F value     Pr(>F)
natural                 17606539    17606539     1 310.97    54.5589 1.397e-12
Dose                  1517357360  1517357360     1 310.97  4701.9667 < 2.2e-16
Substrate              380517369    76103474     5 310.97   235.8284 < 2.2e-16
natural:Dose            15683854    15683854     1 310.97    48.6009 1.886e-11
natural:Substrate        8560933     1712187     5 310.97     5.3057 0.0001077
Dose:Substrate         362539320    72507864     5 310.97   224.6864 < 2.2e-16
natural:Dose:Substrate   8399558     1679912     5 310.97     5.2057 0.0001323

natural                 ***
Dose                    ***
Substrate               ***
natural:Dose            ***
natural:Substrate       ***
Dose:Substrate          ***
natural:Dose:Substrate ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
refitting model(s) with ML (instead of REML)


Data: data
Models:
lmm2: VO2 ~ Dose * Substrate + (1 | pair)
lmm1: VO2 ~ natural * Dose * Substrate + (1 | pair)
     npar    AIC    BIC  logLik -2*log(L)  Chisq Df Pr(>Chisq)
lmm2   14 5632.8 5686.4 -2802.4    5604.8
lmm1   26 5332.5 5432.0 -2640.2    5280.5 324.39 12  < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Likelihood ratio test

Model 1: VO2 ~ natural * Dose * Substrate + (1 | pair)
Model 2: VO2 ~ Dose * Substrate + (1 | pair)
  #Df  LogLik  Df  Chisq Pr(>Chisq)
1  26 -2489.7
2  14 -2724.9 -12 470.45  < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


# R2 for Mixed Models

  Conditional R2: 0.965
     Marginal R2: 0.901


# Intraclass Correlation Coefficient

    Adjusted ICC: 0.644
  Unadjusted ICC: 0.064



Call:
lm(formula = VO2 ~ natural * Dose * Substrate, data = data)

Residuals:
    Min      1Q  Median      3Q     Max
-3254.1  -370.8   186.2   536.4  1727.3

Coefficients:
```
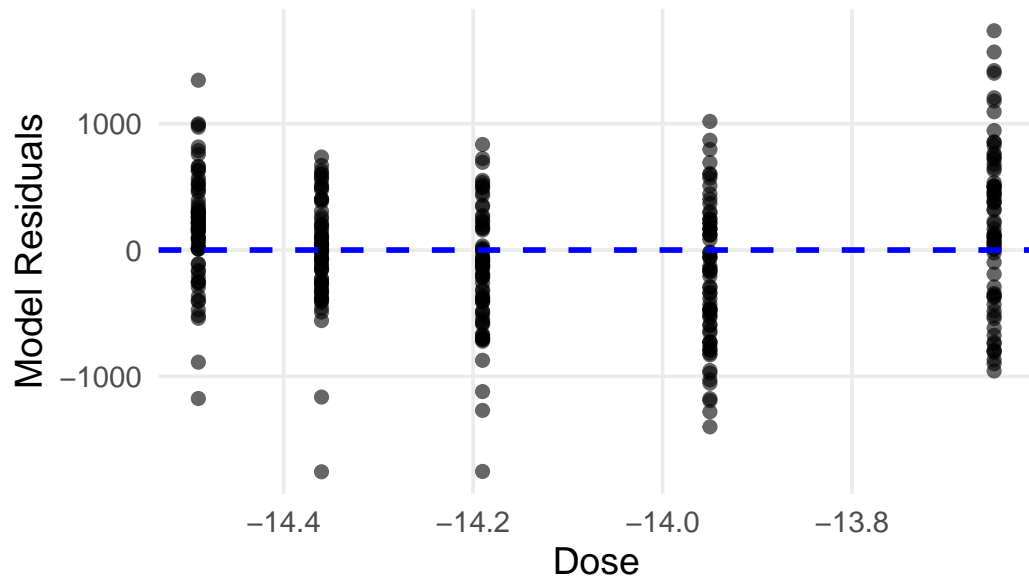
```
                                     Estimate Std. Error t value Pr(>|t|)
(Intercept)                          143278.42    7230.68  19.815  < 2e-16 ***
naturalTransgenic                     22297.32   10225.72   2.181   0.0300 *
Dose                                   9596.47     511.68  18.755  < 2e-16 ***
SubstrateOcM                        -109203.79   10724.83 -10.182  < 2e-16 ***
SubstratePcM                        -112973.48   10724.83 -10.534  < 2e-16 ***
SubstratePM                           -1342.39   10225.72  -0.131   0.8956
SubstratePMOc                         -8026.61   10225.72  -0.785   0.4331
SubstratePMPc                        -54350.93   10225.72  -5.315 2.02e-07 ***
naturalTransgenic:Dose                 1513.62     723.63   2.092   0.0373 *
naturalTransgenic:SubstrateOcM       -17567.40   15167.20  -1.158   0.2476
naturalTransgenic:SubstratePcM       -10213.84   15167.20  -0.673   0.5012
naturalTransgenic:SubstratePM        -11623.31   14461.35  -0.804   0.4221
naturalTransgenic:SubstratePMOc        7909.05   14461.35   0.547   0.5848
naturalTransgenic:SubstratePMPc       26976.20   14461.35   1.865   0.0631 .
Dose:SubstrateOcM                     -7485.71     758.95  -9.863  < 2e-16 ***
Dose:SubstratePcM                     -7737.11     758.95 -10.195  < 2e-16 ***
Dose:SubstratePM                         26.09     723.63   0.036   0.9713
Dose:SubstratePMOc                     -532.46     723.63  -0.736   0.4624
Dose:SubstratePMPc                    -3706.61     723.63  -5.122 5.26e-07 ***
naturalTransgenic:Dose:SubstrateOcM   -1278.14    1073.31  -1.191   0.2346
naturalTransgenic:Dose:SubstratePcM    -784.42    1073.31  -0.731   0.4654
naturalTransgenic:Dose:SubstratePM     -790.30    1023.37  -0.772   0.4405
naturalTransgenic:Dose:SubstratePMOc    550.68    1023.37   0.538   0.5909
naturalTransgenic:Dose:SubstratePMPc   1854.02    1023.37   1.812   0.0710 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 839.7 on 316 degrees of freedom
  (20 observations deleted due to missingness)
Multiple R-squared:  0.9233,    Adjusted R-squared:  0.9177
F-statistic: 165.4 on 23 and 316 DF,  p-value: < 2.2e-16


Warning in modelUpdate(objects[[i - 1]], objects[[i]]): original model was of
class "lmerModLmerTest", updated model is of class "lm"


Likelihood ratio test

Model 1: VO2 ~ natural * Dose * Substrate + (1 | pair)
Model 2: VO2 ~ natural * Dose * Substrate
  #Df  LogLik Df Chisq Pr(>Chisq)
1  26 -2489.7
```
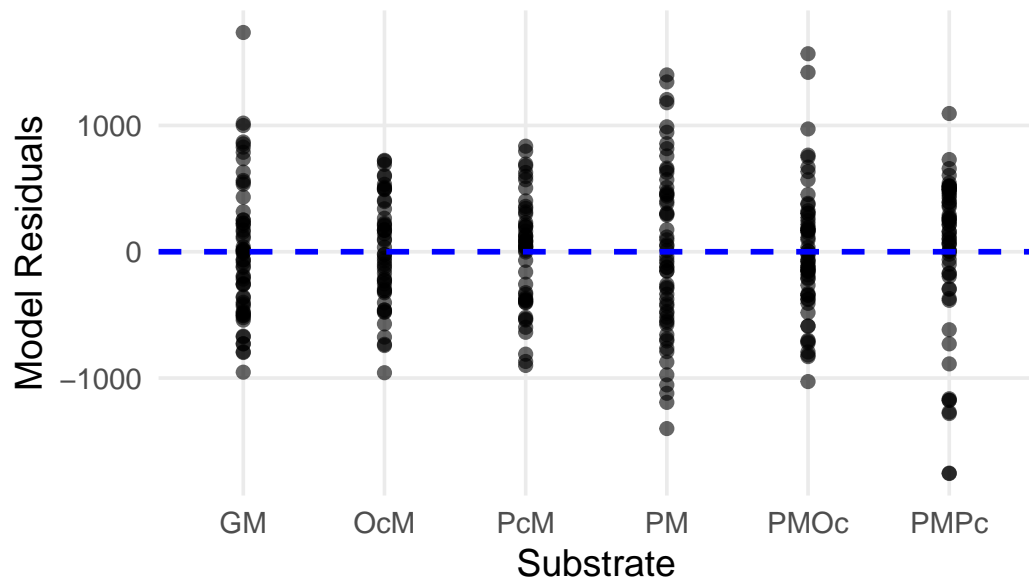
```
2  25 -2759.2 -1 539.1  < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```
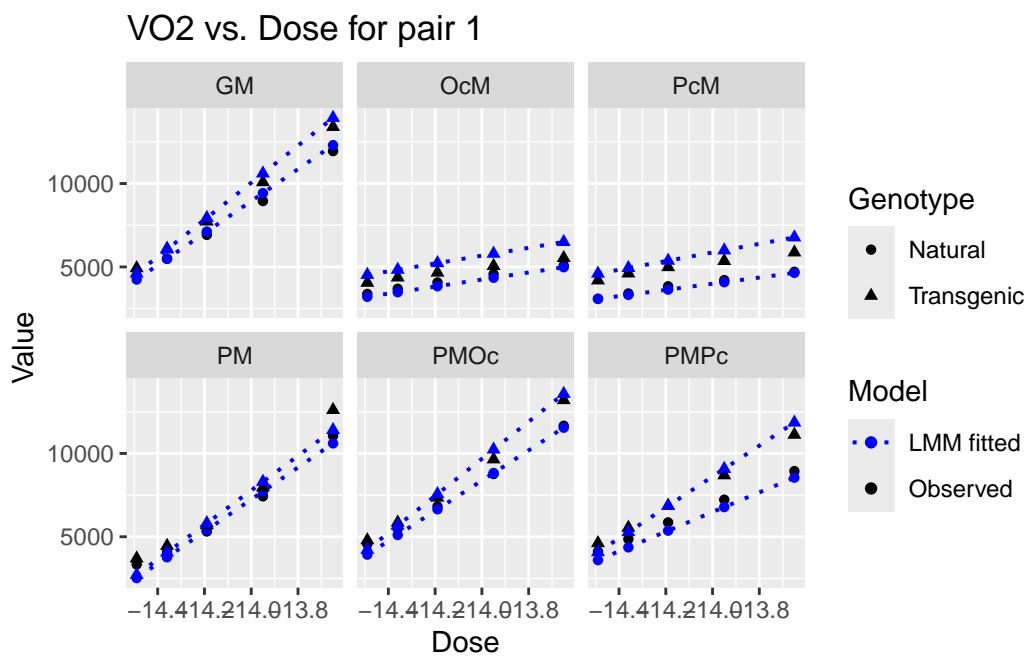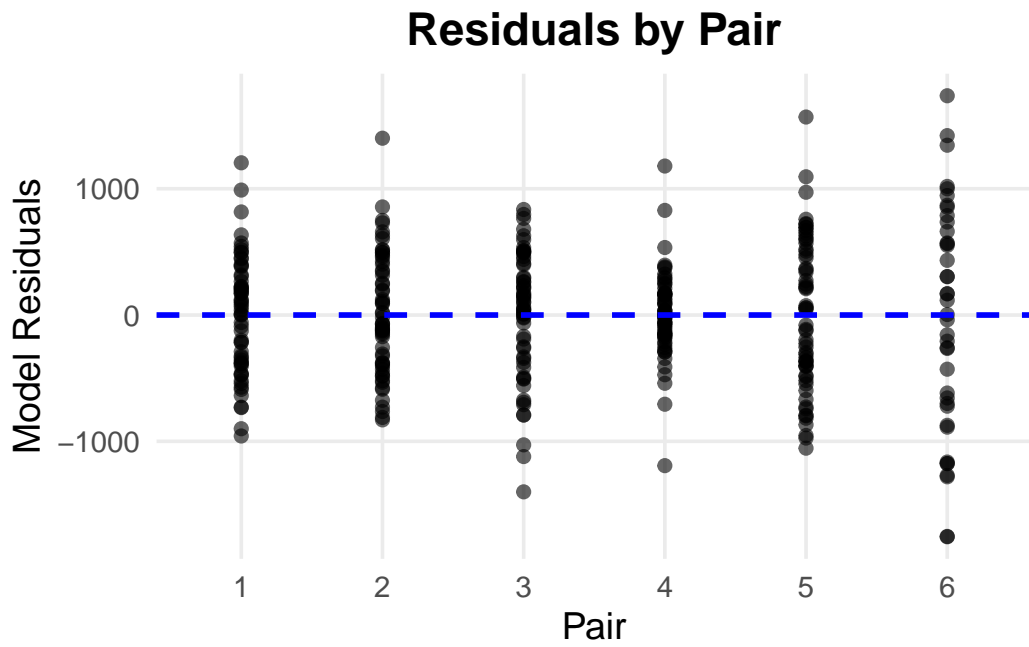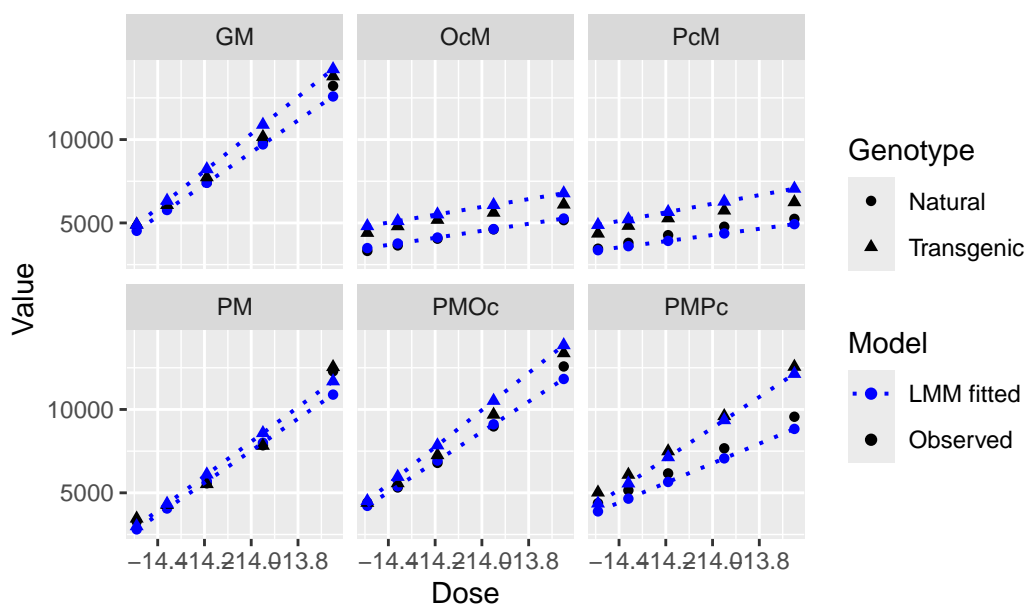
Residuals

Random Effects

Residuals by Genotype
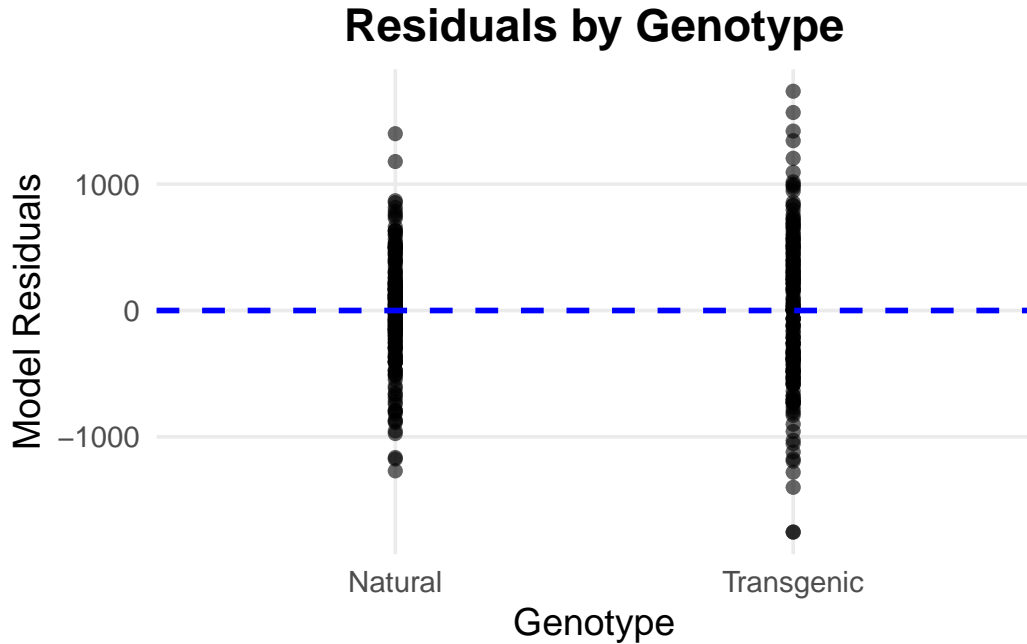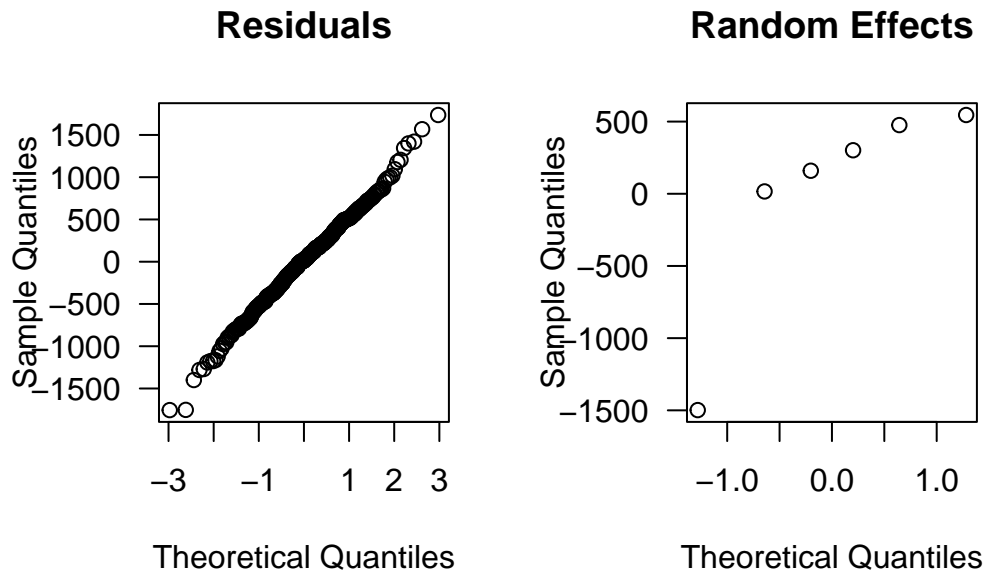
Residuals by Dose

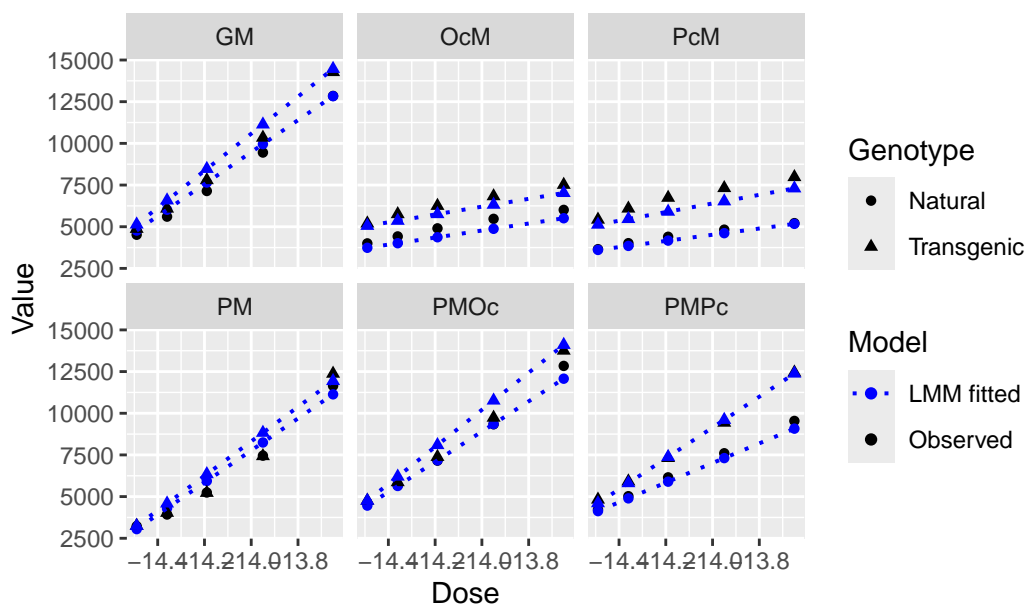# Residuals by Substrate

Residuals by Pair

VO2 vs. Dose for pair 1

VO2 vs. Dose for pair 2
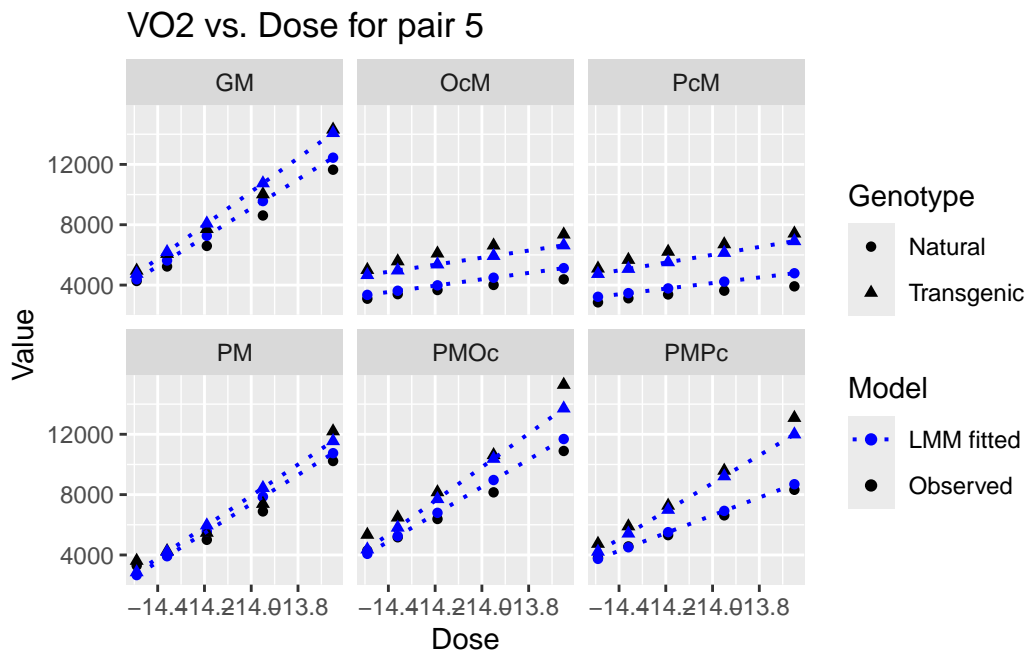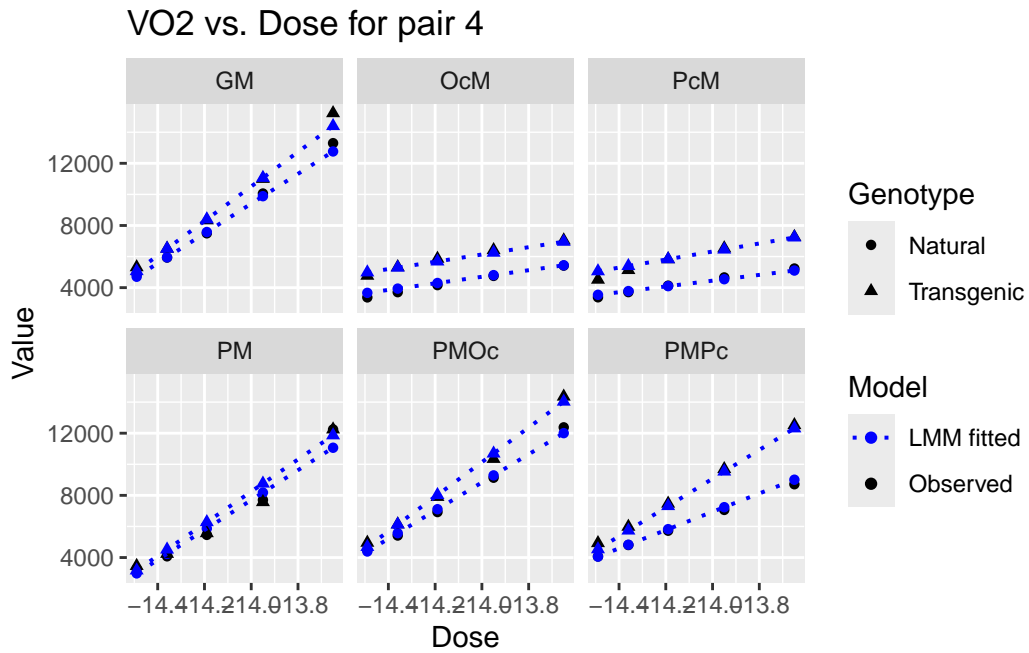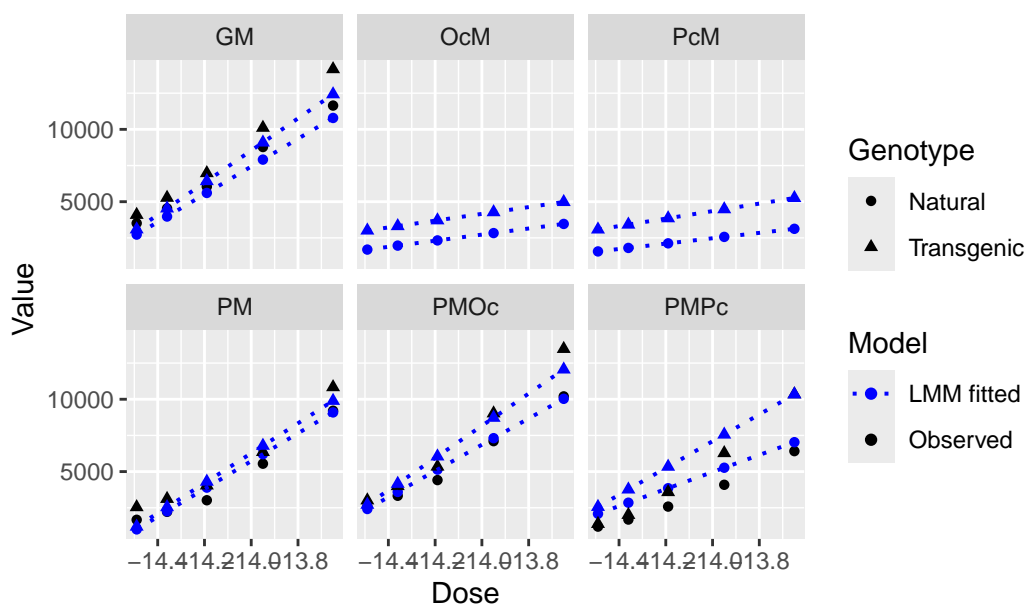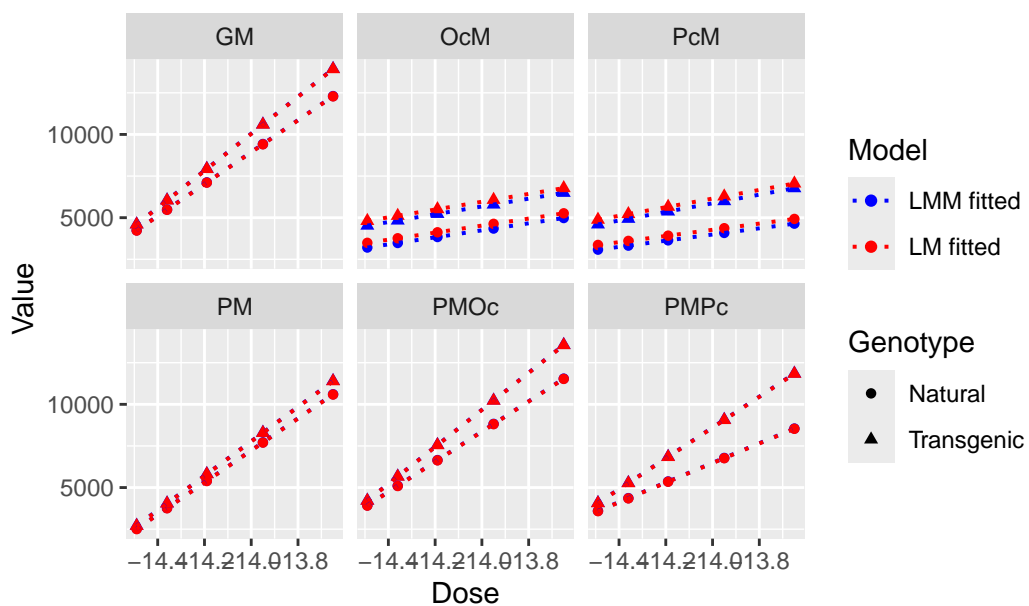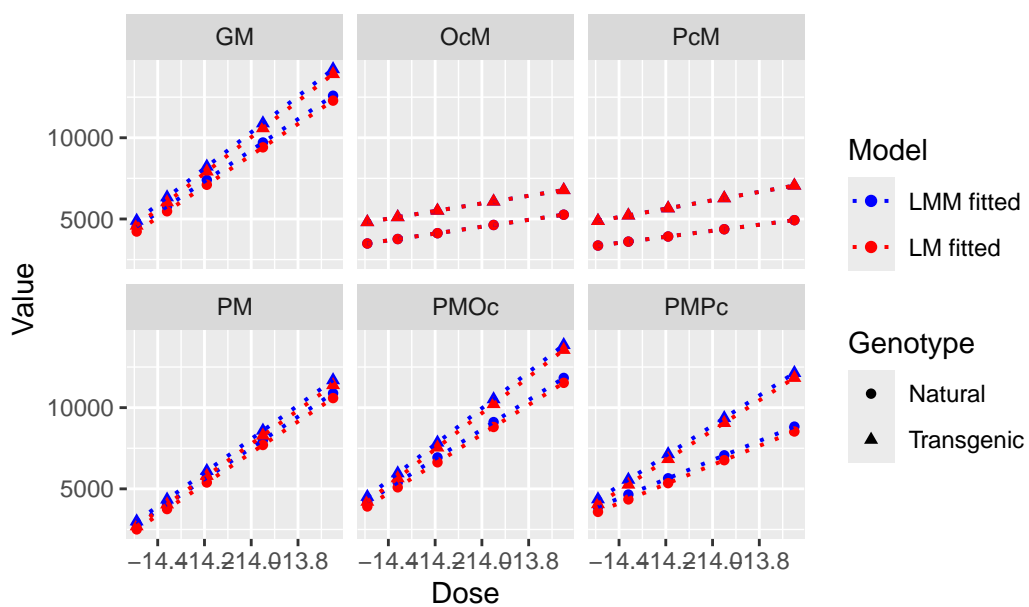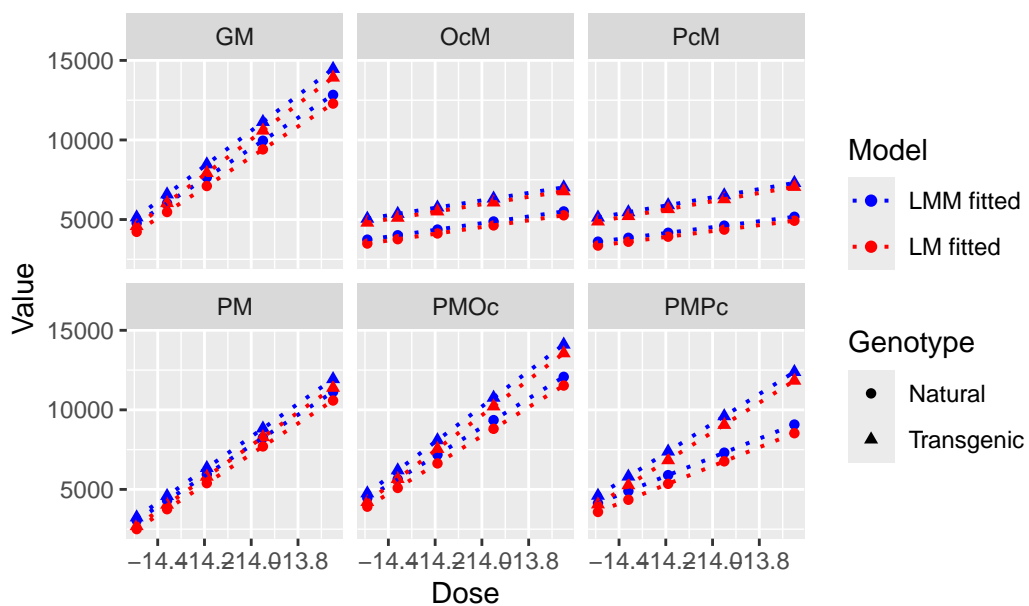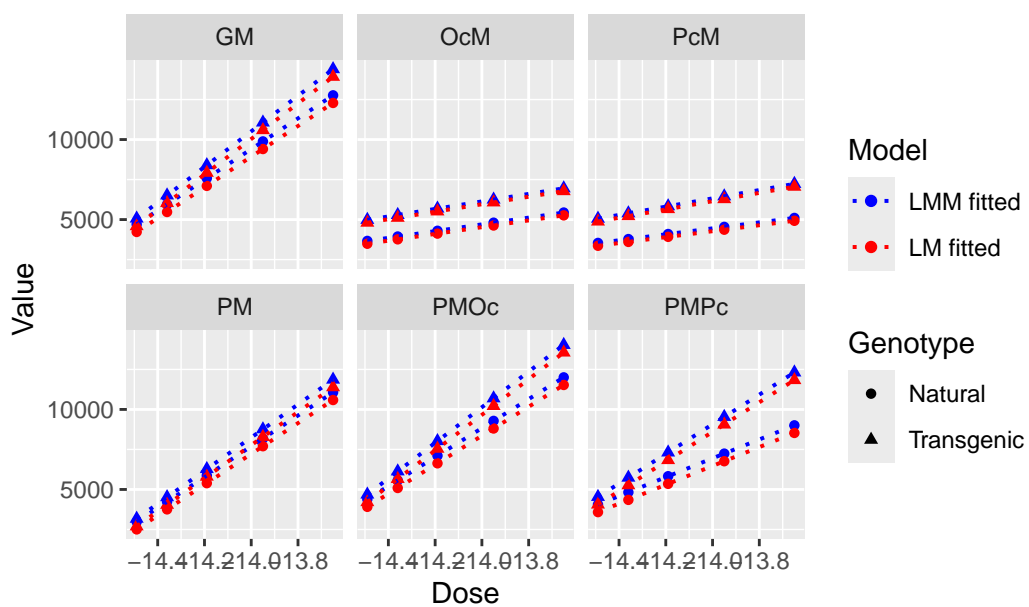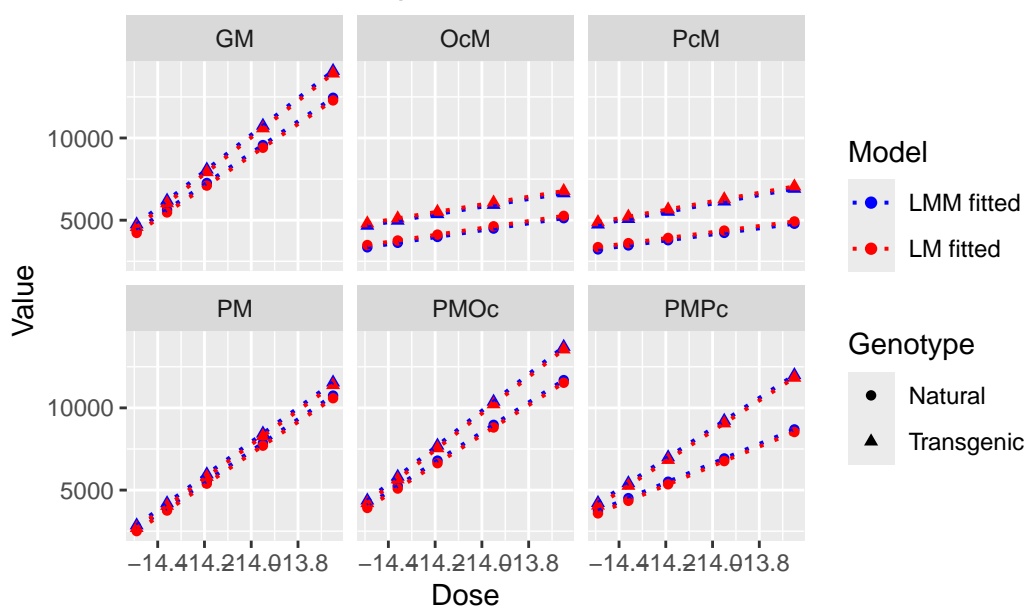

VO2 vs. Dose for pair 3

VO2 vs. Dose for pair 4

VO2 vs. Dose for pair 5

Warning: Removed 20 rows containing missing values or values outside the scale range
(`geom_point()`).

VO2 vs. Dose for pair 2
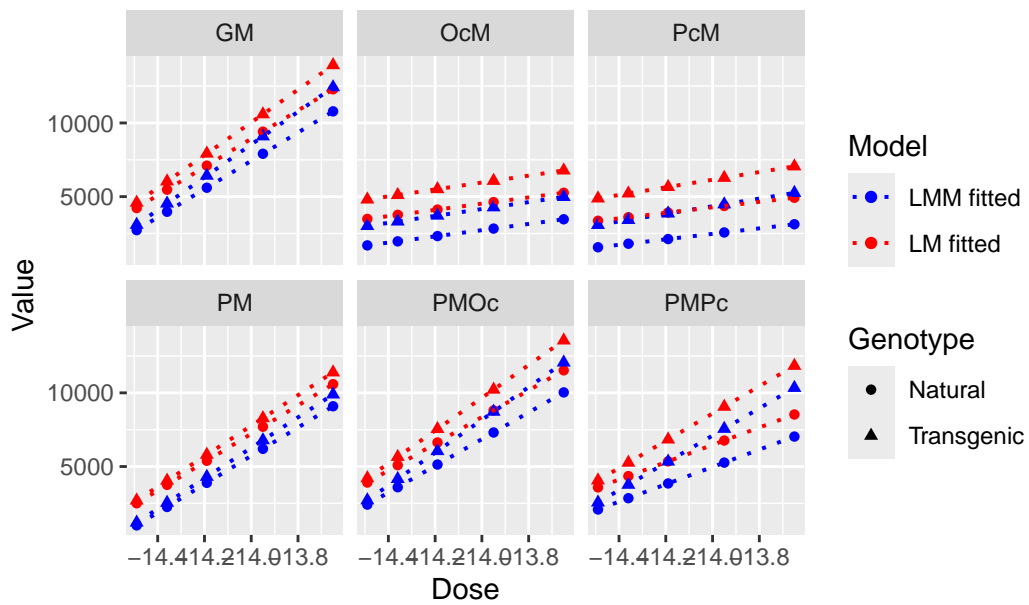

VO2 vs. Dose for pair 3
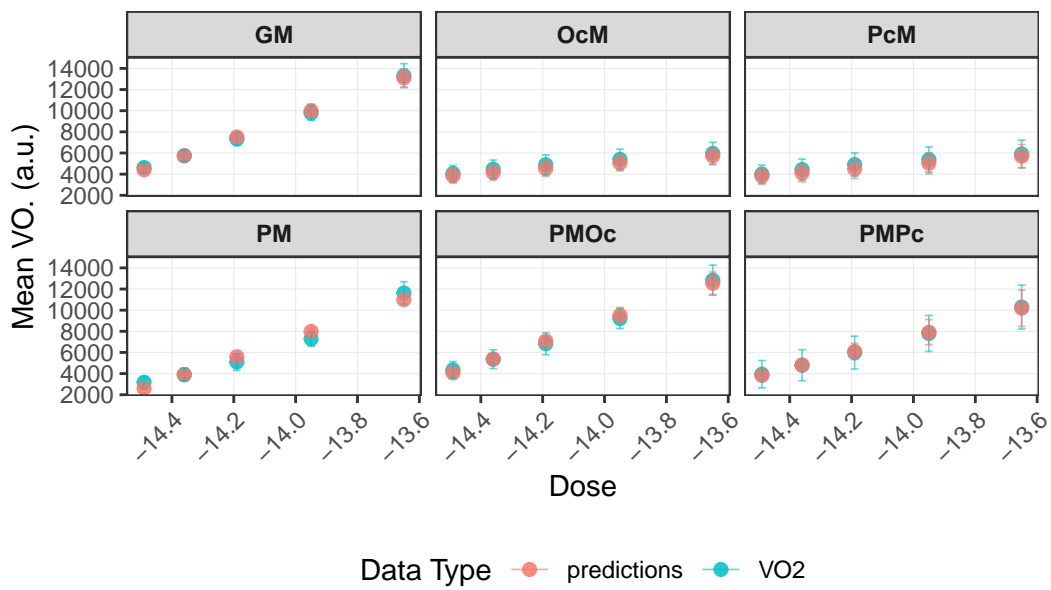
VO2 vs. Dose for pair 4

VO2 vs. Dose for pair 5

VO2 vs. Dose for pair 6

Dose vs. Mean VO.

Predictions vs. Observed VO.