

# Report: Hybrid vigor in response to Eimeria in the HMHZ

*Alice*

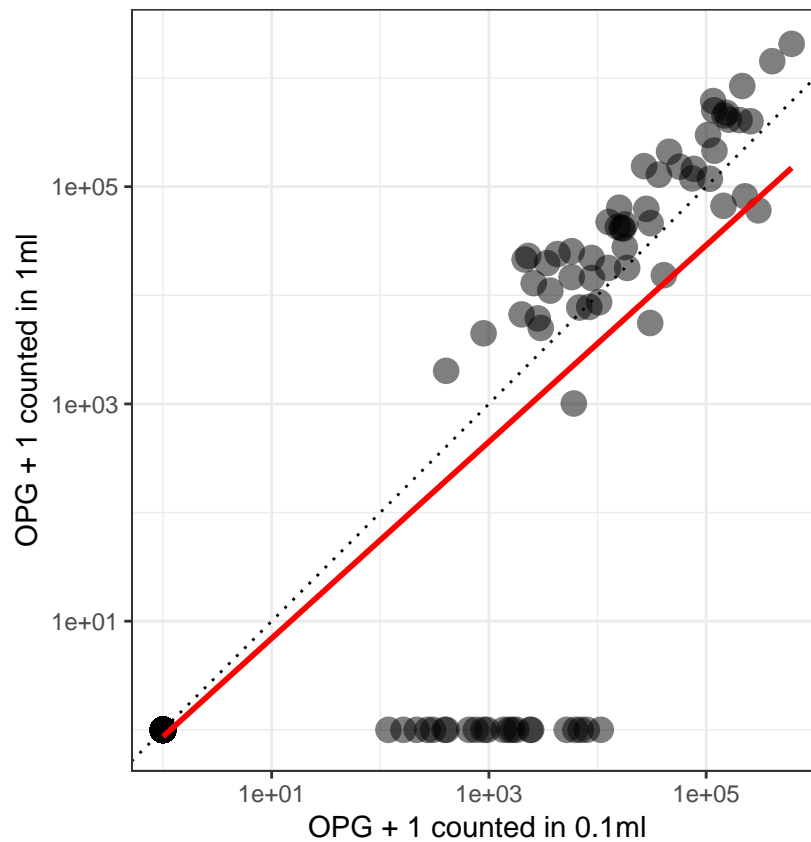
*15 October 2018*

## Contents

<b>Eimeria detection oocysts flotation</b>	<b>2</b>
Improving Eimeria oocysts detection . . . . .	2
OPG that we keep . . . . .	2
<b>Eimeria detection PCR</b>	<b>3</b>
<b>Eimeria detection qPCR</b>	<b>4</b>
<b>General stats on sampling</b>	<b>7</b>
<b>General informations on HMHZ</b>	<b>7</b>
<b>Prevalence of our 3 different methods</b>	<b>8</b>
Prevalence tables . . . . .	8
OPG-PCR . . . . .	9
OPG-qPCR . . . . .	10
OPG-qPCR-PCR . . . . .	10
<b>Testing hybrid vigor along HMHZ</b>	<b>10</b>

# Eimeria detection oocysts flotation

## Improving Eimeria oocysts detection



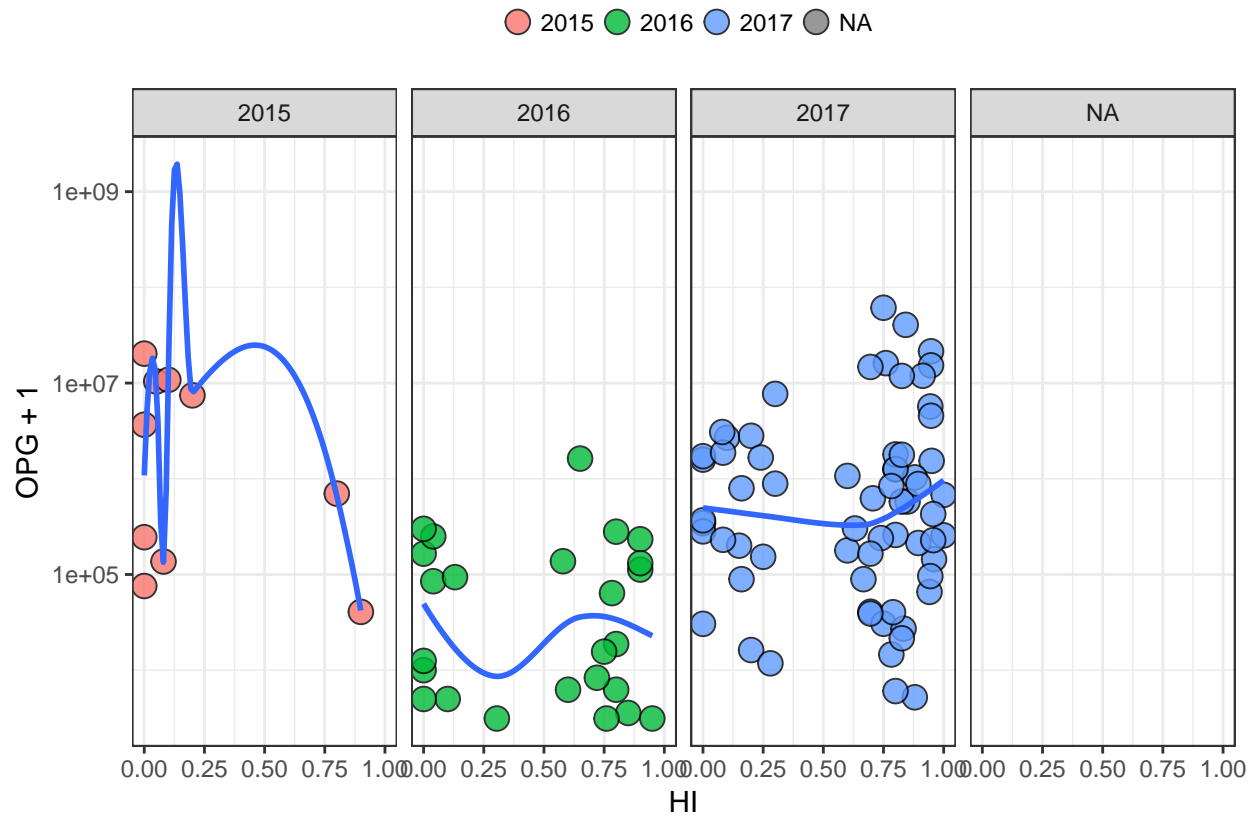
22 new samples were detected while diluting by 0.1mL PBS instead of 1mL before counting in Neubauer chamber.

Adjusted R-squared = 0.81 represents the amount of variation in y explained by x.

## OPG that we keep

Number of Mus musculus caught with OPG values: 484

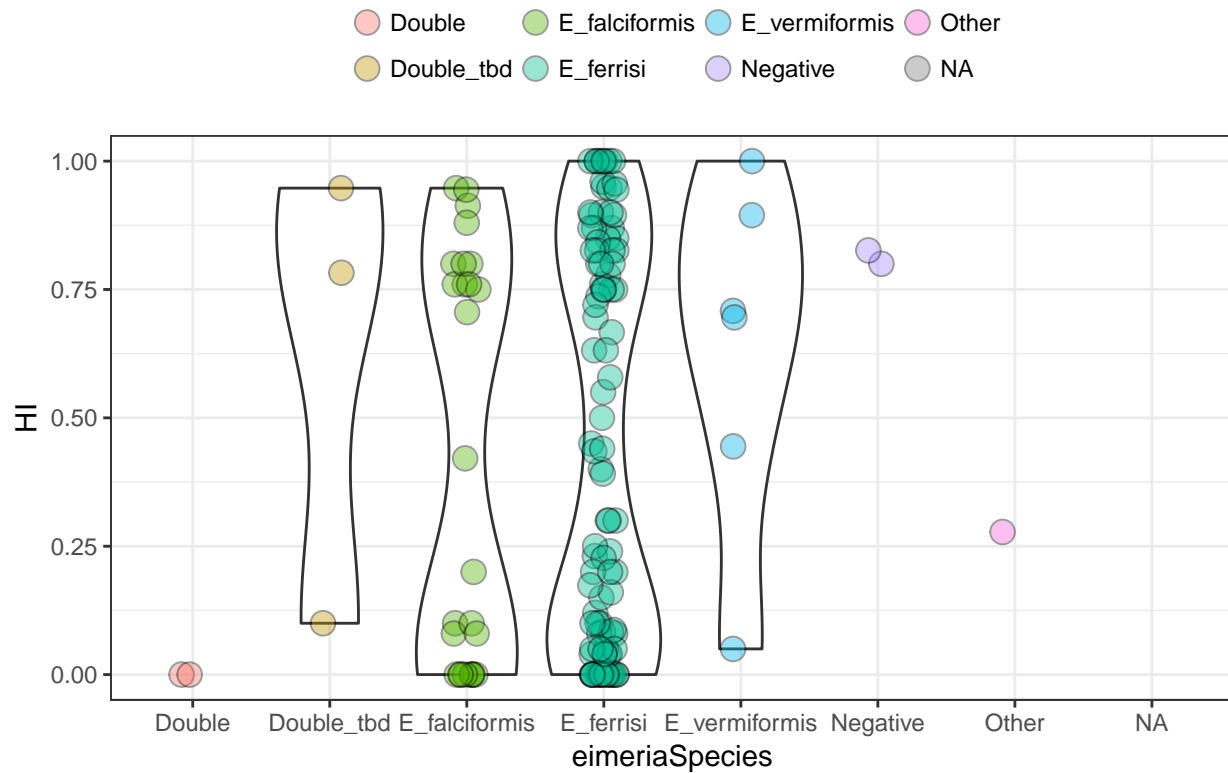
```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```



## Eimeria detection PCR

PCR positive = one of the 3 other markers than AP5 sequenced (Ap5 was used for detection only, the other markers for confirmation)

## Violin plots on PCR data

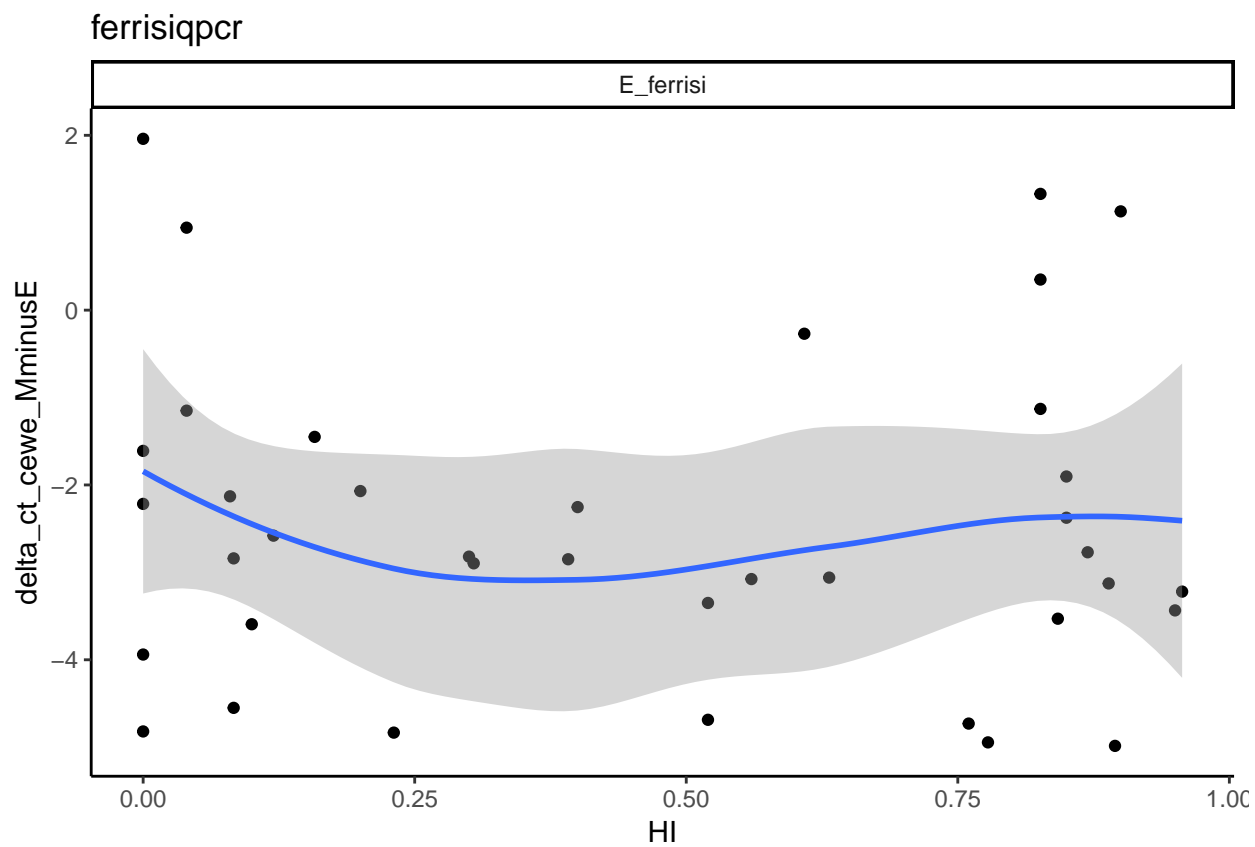


PCR positive = one of the 3 markers 18S, COI or ORF470) gave a sequence. Number of *Mus musculus* caught with PCR performed: 1165

## Eimeria detection qPCR

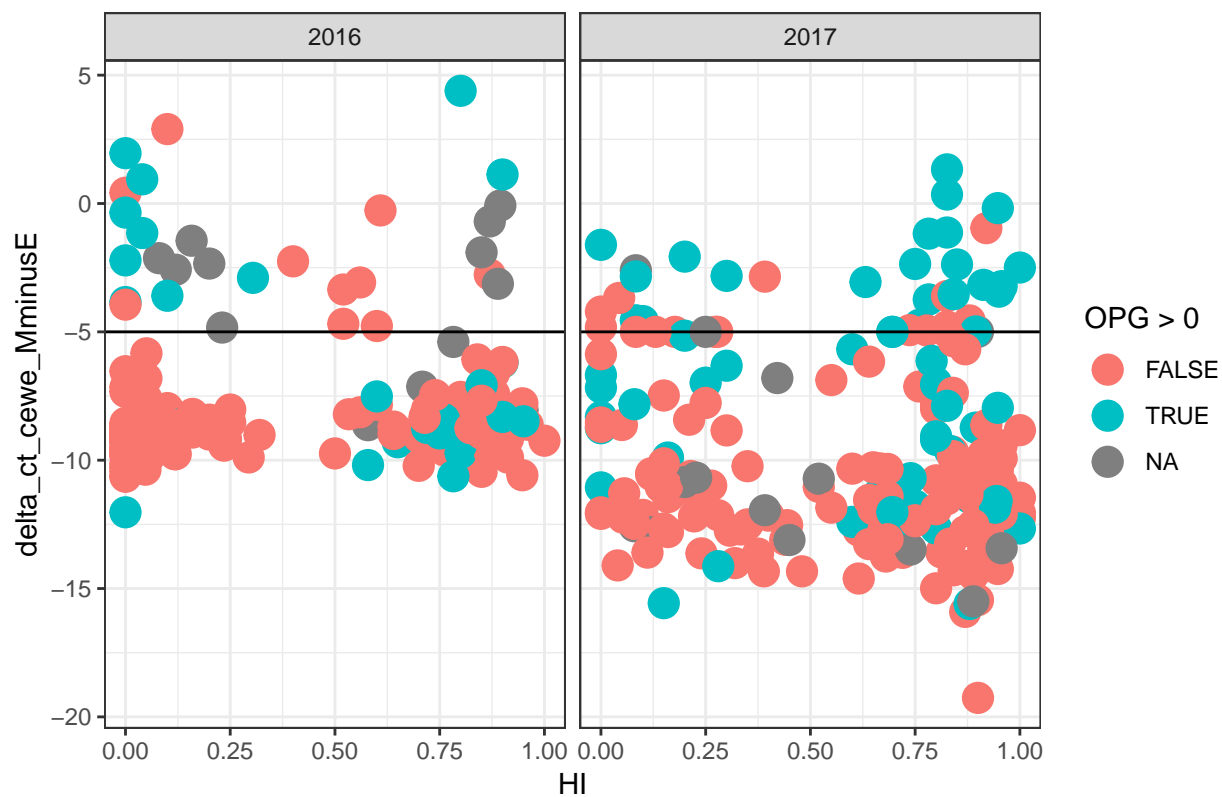
We keep only the values for mice having been tested for BOTH ileum and cecum!

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```

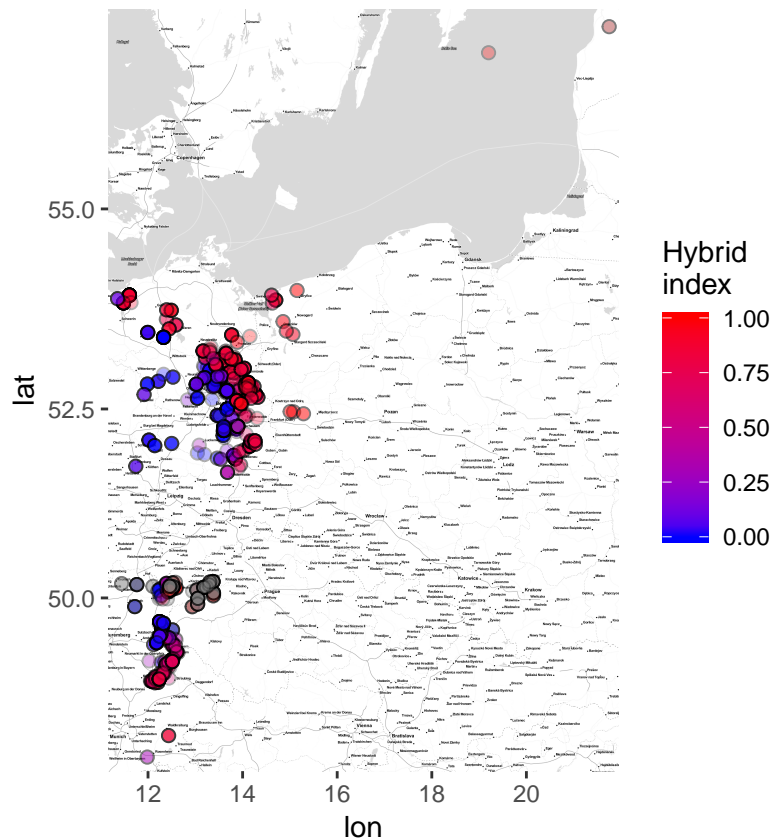


## Warning: Removed 107 rows containing missing values (geom\_point).

Remark of J. Wolinska: some individuals here HAVE qPCR value, but no  $\alpha$



## General stats on sampling



- Some information regarding latitude and longitude are missing for the following mice:

AA\_0080c, AA\_0080i, AA\_0100c, SK\_2903, SK\_2904, SK\_2958, SK\_2959, SK\_2960, SK\_2961, SK\_3174, SK\_3333, SK\_3417, SK\_3461, SK\_3462, SK\_3469, SK\_3470, SK\_3471

- We still miss info (HI) on the following mice (ask Jarda):

AA\_0080c, AA\_0080i, AA\_0100c, AA\_0411, AA\_0412, AA\_0420, AA\_0464, SK\_2668, SK\_2669, SK\_2671, SK\_2674, SK\_2675, SK\_2676, SK\_2677, SK\_2678, SK\_2681, SK\_2682, SK\_2684, SK\_2685, SK\_2687, SK\_2688, SK\_2690, SK\_2692, SK\_2693, SK\_2695, SK\_2696, SK\_2699, SK\_2700, SK\_2701, SK\_2702, SK\_2703, SK\_2704, SK\_2705, SK\_2710, SK\_2713, SK\_2715, SK\_2724, SK\_2727, SK\_2729, SK\_2733, SK\_2734, SK\_2736, SK\_2737, SK\_2738, SK\_2739, SK\_2745, SK\_2750, SK\_2751, SK\_2752, SK\_2754, SK\_2755, SK\_2756, SK\_2758, SK\_2759, SK\_2760, SK\_2761, SK\_2775, SK\_2778, SK\_2780, SK\_2782, SK\_2789, SK\_2792, SK\_2793, SK\_2794, SK\_2795, SK\_2798, SK\_2799, SK\_2800, SK\_2801, SK\_2802, SK\_2803, SK\_2804, SK\_2805, SK\_2851, SK\_2852, SK\_2853, SK\_2854, SK\_2855, SK\_2856, SK\_2857, SK\_2858, SK\_2859, SK\_2860, SK\_2861, SK\_2862, SK\_2863, SK\_2864, SK\_2865, SK\_2866, SK\_2868, SK\_2869, SK\_2870, SK\_2871, SK\_2873, SK\_2874, SK\_2875, SK\_2876, SK\_2877, SK\_2878, SK\_2879, SK\_2880, SK\_2881, SK\_2884, SK\_2885, SK\_2886, SK\_2887, SK\_2888, SK\_2889, SK\_2958, SK\_2959, SK\_2960, SK\_2961, SK\_3174, SK\_3333, SK\_3417, SK\_3461, SK\_3462, SK\_3469, SK\_3470, SK\_3471

## General informations on HMHz

- 1339 mice were captured over three years, from 285 farms

- On average, NA mice were caught per farm (95% CI NA)
- **Hybrid indexes** were calculated as ratio of M.m.d/M.m.m alleles (between 1 and 14, on average 12 loci)

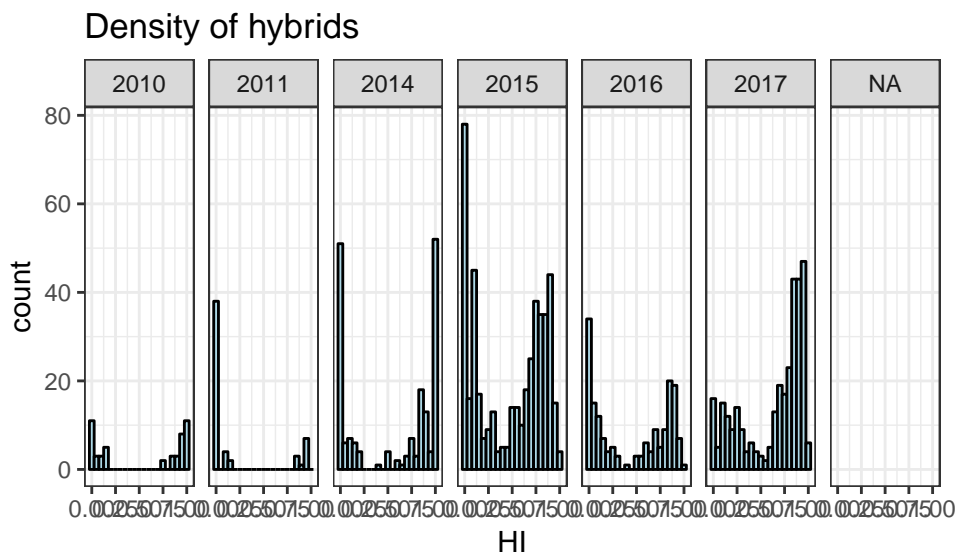


Figure 1: Number of animals caught along the hybrid index

## Prevalence of our 3 different methods

### Prevalence tables

Table 1: Prevalence of Eimeria per year, based on oocyst flotation

	2010	2011	2014	2015	2016	2017
FALSE	0	0	0	92.0	126.00	165.00
TRUE	0	0	0	10.0	25.00	65.00
total	0	0	0	102.0	151.00	230.00
prevalence(%)	NaN	NaN	NaN	9.8	16.56	28.26

Table 2: Prevalence of Eimeria per year, based on PCR detection. A mouse was considered infected by Eimeria if one of the 3 markers (COI, 18S or ORF470) gave a sequence

	2010	2011	2014	2015	2016	2017
FALSE	0	0	241.00	417.00	149.00	207.00
TRUE	0	0	40.00	27.00	21.00	62.00
total	0	0	281.00	444.00	170.00	269.00
prevalence(%)	NaN	NaN	14.23	6.08	12.35	23.05



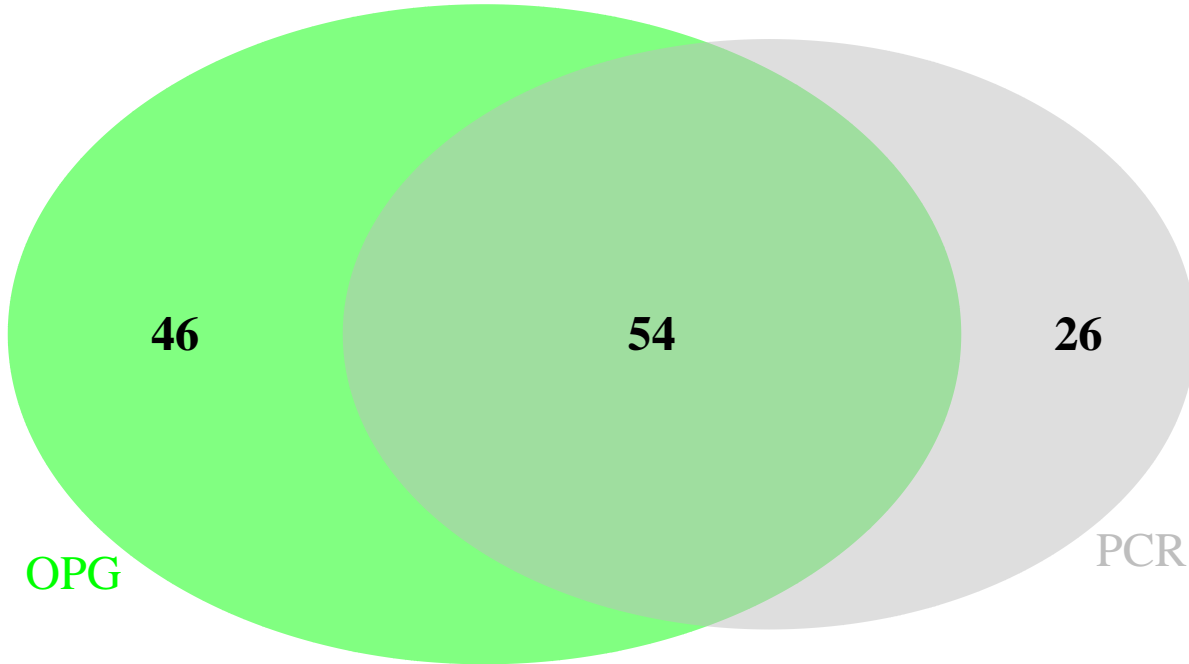
Table 3: Prevalence of Eimeria per year, based on qPCR in cecum

	2010	2011	2014	2015	2016	2017
FALSE	0	0	0	0	136.00	187
TRUE	0	0	0	0	29.00	33
total	0	0	0	0	165.00	220
prevalence(%)	NaN	NaN	NaN	NaN	17.58	15

Table 4: Prevalence of Eimeria per year, based on all detections methods. A mouse was considered infected by Eimeria if one of the 3 markers (COI, 18S or ORF470) gave a sequence, OR if it had a positive count of oocysts in its feces, OR if it was qPCR positive in cecum tissue

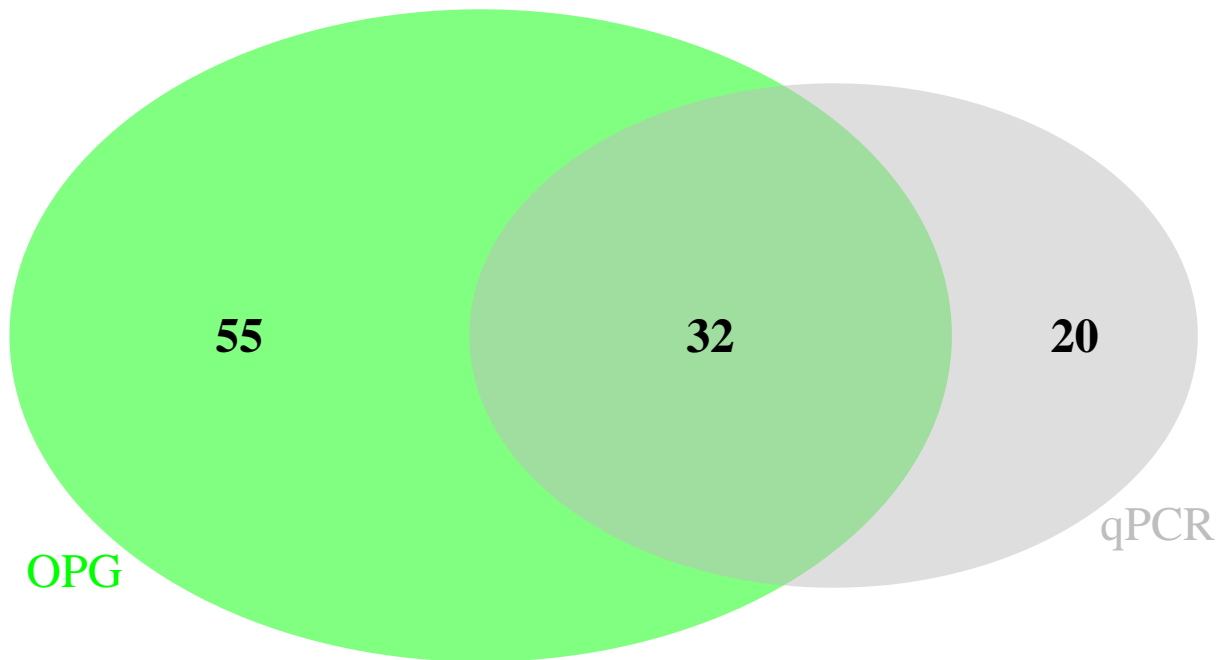
	2010	2011	2014	2015	2016	2017
negative	49	55	243.00	428.00	123.00	227.00
positive	0	0	40.00	32.00	47.00	93.00
total	49	55	283.00	460.00	170.00	320.00
prevalence(%)	0	0	14.13	6.96	27.65	29.06

#### OPG-PCR



## (polygon[GRID.polygon.829], polygon[GRID.polygon.830], polygon[GRID.polygon.831], polygon[GRID.polygon.832])

## OPG-qPCR



```
## (polygon[GRID.polygon.838], polygon[GRID.polygon.839], polygon[GRID.polygon.840], polygon[GRID.polygon.841])
```

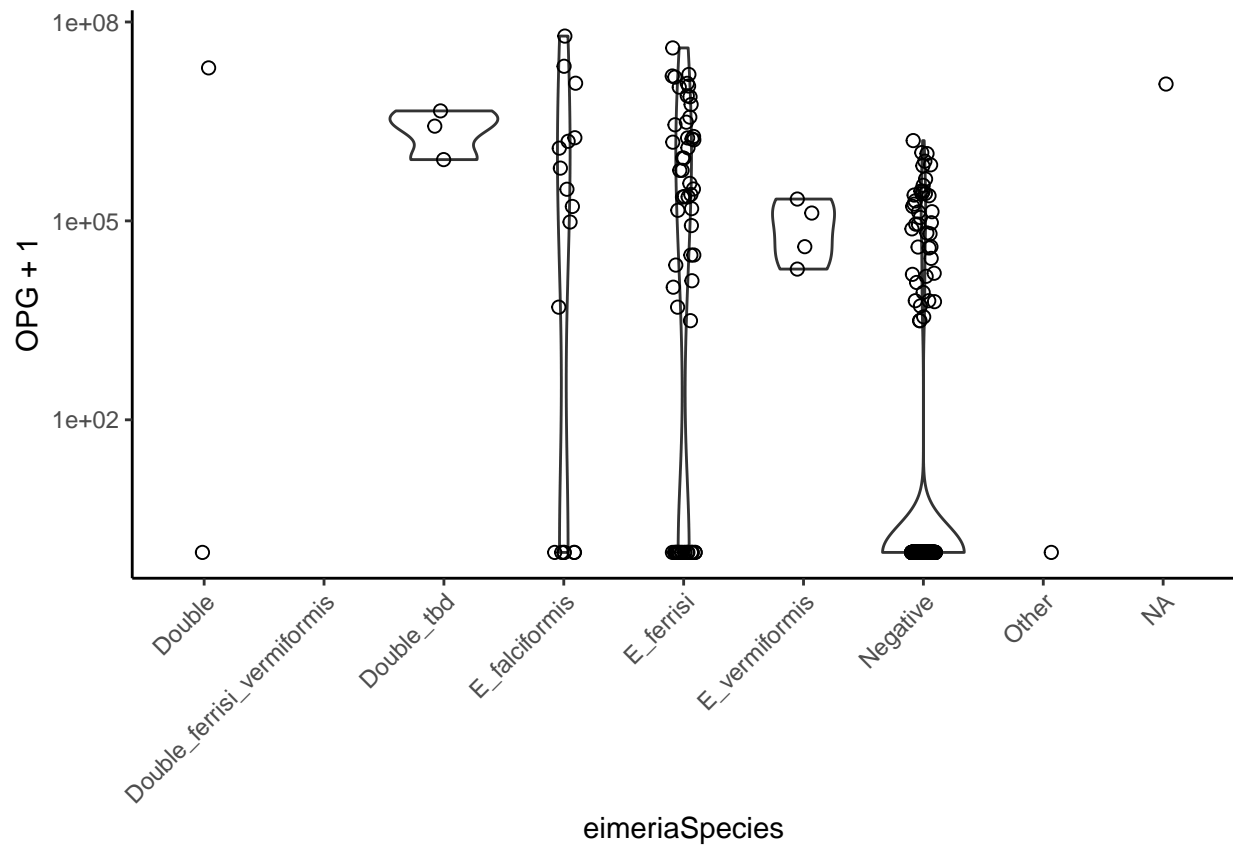
## OPG-qPCR-PCR

### Testing hybrid vigor along HMMZ

```
ggplot(miceTable, aes(x = eimeriaSpecies, y = OPG +1)) +  
  geom_violin() +  
  geom_jitter(width = .1, pch = 21, size = 2) +  
  scale_y_log10() +  
  theme_classic() +  
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```

```
## Warning: Removed 855 rows containing non-finite values (stat_ydensity).
```

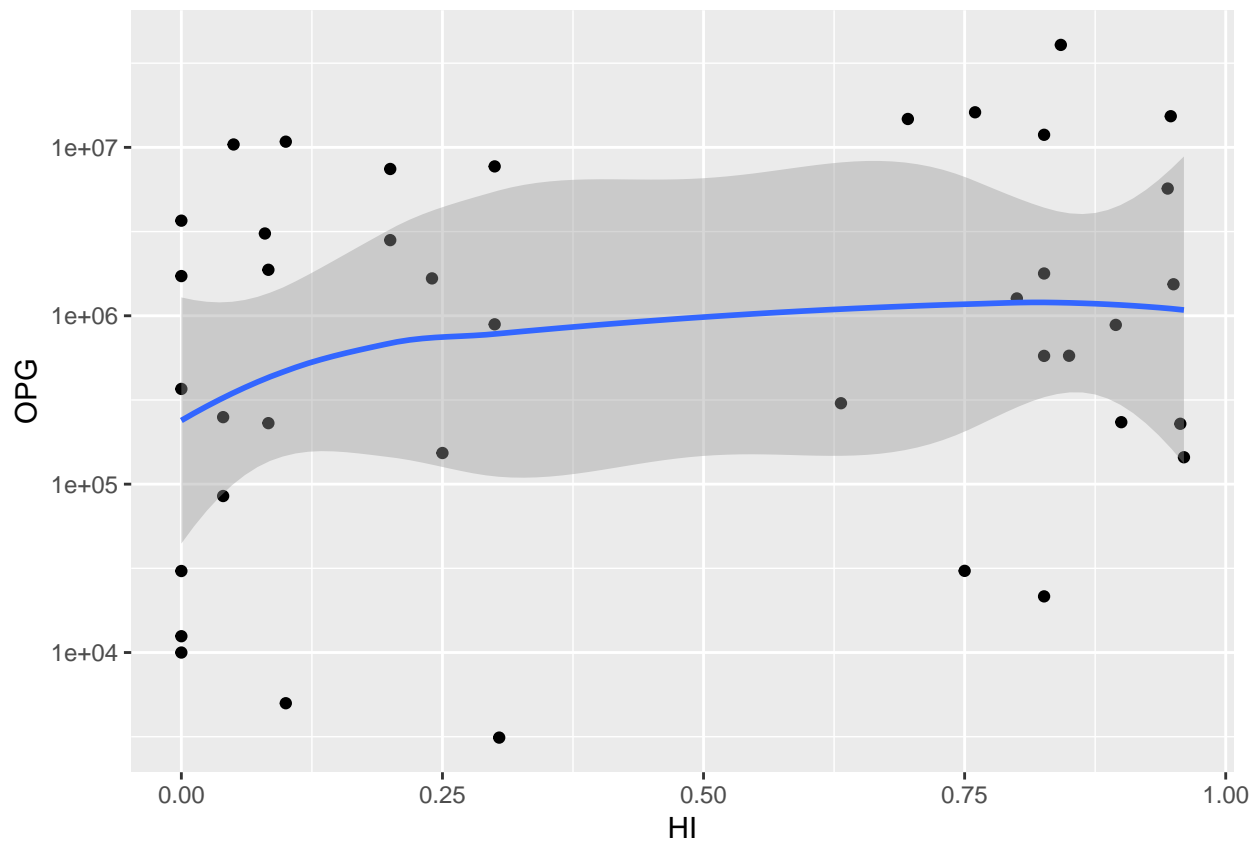
```
## Warning: Removed 855 rows containing missing values (geom_point).
```



```
ferrisiDFfplot <- miceTable[!is.na(miceTable$OPG) &
                             miceTable$OPG > 0 &
                             !is.na(miceTable$eimeriaSpecies) &
                             miceTable$eimeriaSpecies == "E_ferrii",]

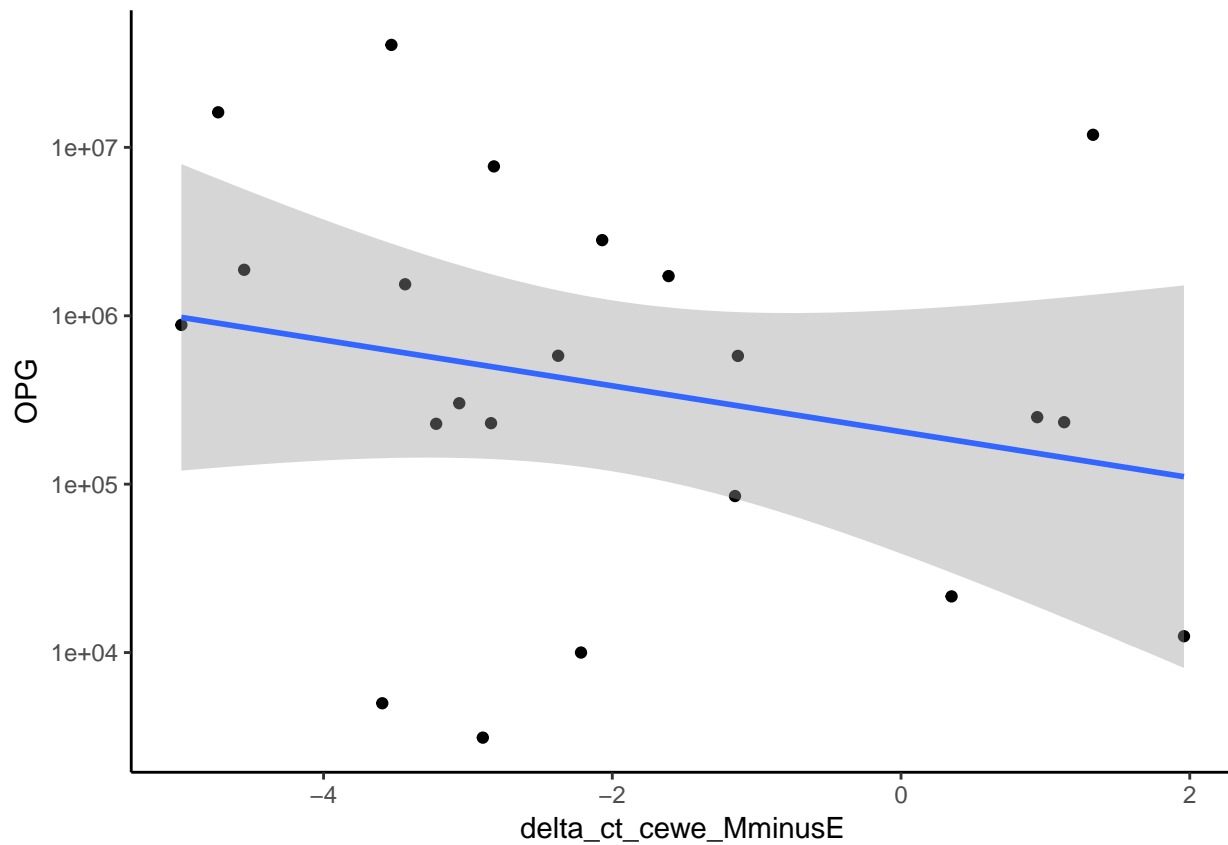
ggplot(ferrisiDFfplot, aes(x = HI, y = OPG)) +
  geom_point() +
  scale_y_log10() +
  geom_smooth()
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```



```
## qPCR by OPG
ferrisiDFflot_n_qpcr <- miceTable[!is.na(miceTable$OPG) &
  miceTable$OPG > 0 &
  !is.na(miceTable$delta_ct_cewe_MminusE) &
  miceTable$delta_ct_cewe_MminusE > -5 &
  !is.na(miceTable$eimeriaSpecies) &
  miceTable$eimeriaSpecies == "E_ferrisi",]

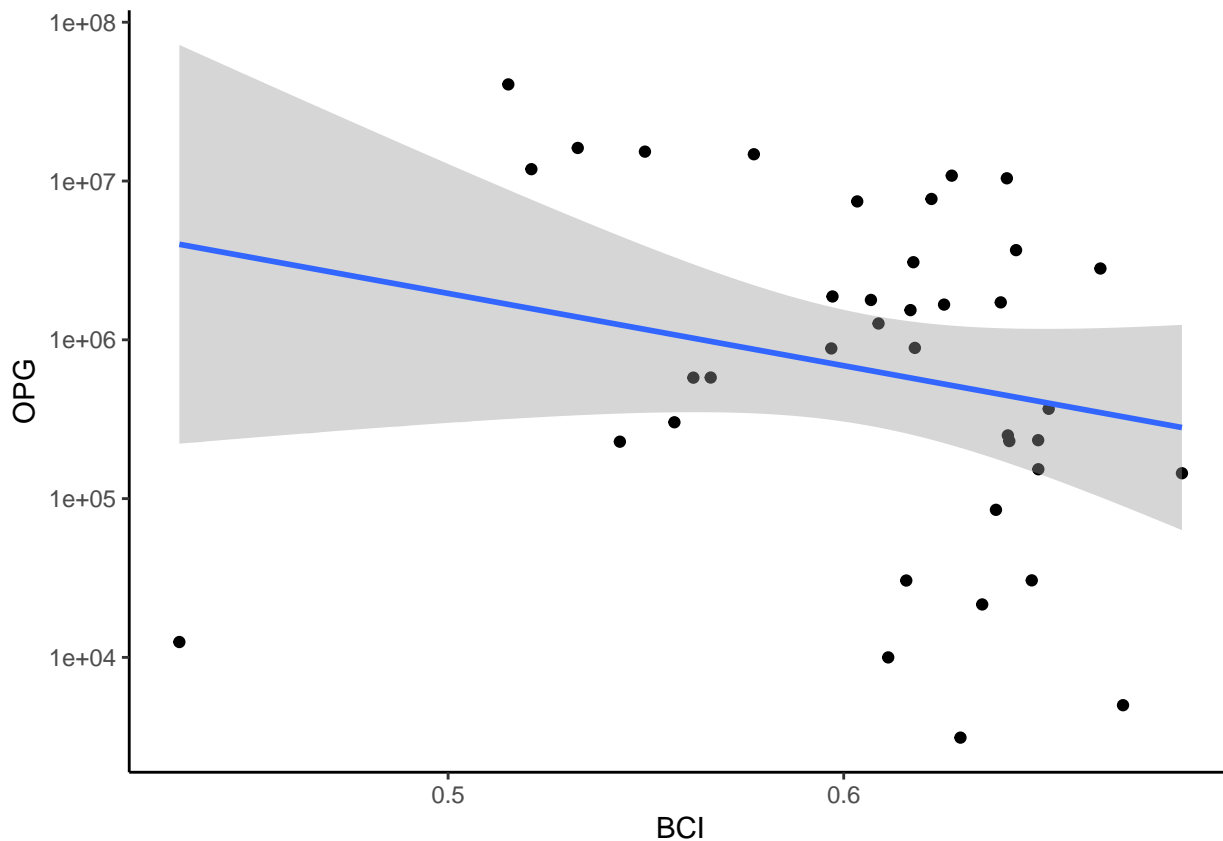
ggplot(ferrisiDFflot_n_qpcr, aes(x = delta_ct_cewe_MminusE, y = OPG)) +
  geom_point() +
  scale_y_log10() +
  geom_smooth(method = "lm") +
  theme_classic()
```



```
## BCI vs OPG
ggplot(ferrisiDFflot, aes(x = BCI, y = OPG)) +
  geom_point() +
  scale_y_log10() +
  geom_smooth(method = "lm") +
  theme_classic()
```

```
## Warning: Removed 1 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 1 rows containing missing values (geom_point).
```

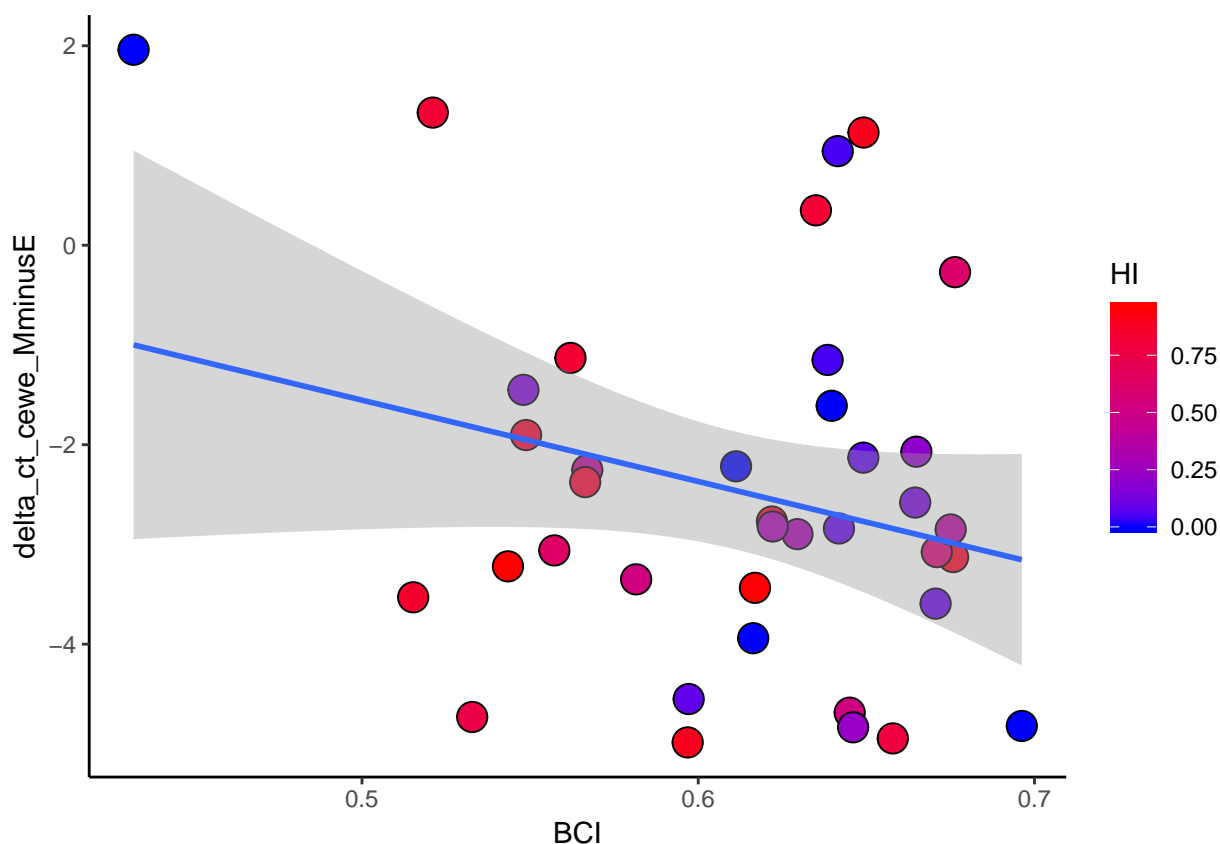


```
summary(lm(ferrisiDFflot$BCI ~ ferrisiDFflot$OPG))
```

```
##
## Call:
## lm(formula = ferrisiDFflot$BCI ~ ferrisiDFflot$OPG)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.186174 -0.013713  0.008491  0.028530  0.067600
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    6.183e-01  8.720e-03  70.908 < 2e-16 ***
## ferrisiDFflot$OPG -2.764e-09  9.995e-10  -2.765  0.00891 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.04712 on 36 degrees of freedom
## (1 observation deleted due to missingness)
## Multiple R-squared:  0.1752, Adjusted R-squared:  0.1523
## F-statistic: 7.648 on 1 and 36 DF, p-value: 0.00891
```

```
## BCI vs qPCR
ggplot(ferrisiDFqpcr,
  aes(x = BCI, y = delta_ct_cewe_MminusE)) +
  geom_point(aes(fill = HI), pch = 21, size = 5) +
  scale_fill_gradient(low = "blue", high = "red") +
  geom_smooth(method = "lm") +
```

```
theme_classic()
```

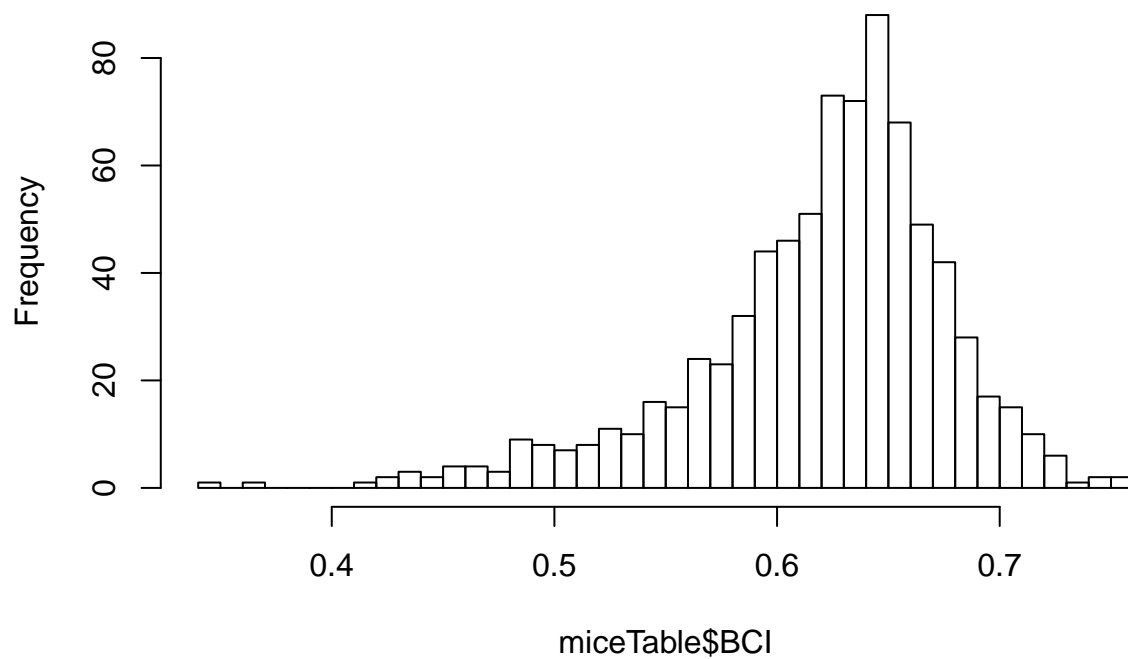


```
summary(lm(ferrisiDFqpcr$BCI ~ ferrisiDFqpcr$delta_ct_cewe_MminusE))
```

```
##
## Call:
## lm(formula = ferrisiDFqpcr$BCI ~ ferrisiDFqpcr$delta_ct_cewe_MminusE)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.14333 -0.04150  0.01148  0.04930  0.08291
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.591278   0.015494  38.162  <2e-16
## ferrisiDFqpcr$delta_ct_cewe_MminusE -0.008099   0.005076  -1.596   0.119
##
## (Intercept)                ***
## ferrisiDFqpcr$delta_ct_cewe_MminusE
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.05654 on 36 degrees of freedom
## Multiple R-squared:  0.06604,    Adjusted R-squared:  0.0401
## F-statistic: 2.546 on 1 and 36 DF,  p-value: 0.1193
```

```
hist(miceTable$BCI, breaks = 50)
```

## Histogram of miceTable\$BCI



```
## BCI vs qPCR
ferrisiDFqpcr$delta_ct_cewe_MminusE

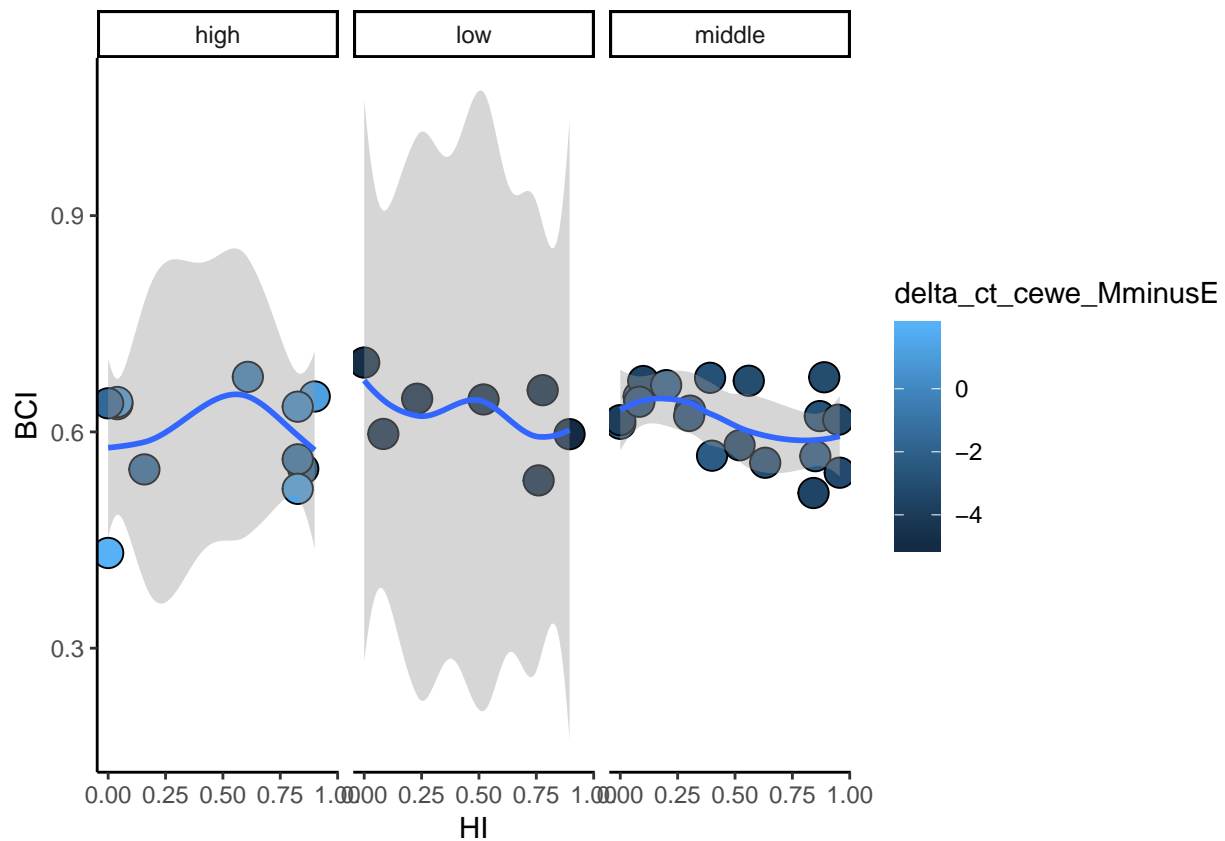
## [1] -0.2700000 -2.1300000 -2.2166667 -2.5800000 -3.5933333 -1.1500000
## [7] -3.9400000  0.9433333  1.9600000 -1.4500000 -2.7700000 -3.1266667
## [13] -1.9033333 -2.2533333 -4.6866667 -3.3500000 -4.8333333  1.1300000
## [19] -2.8966667 -3.0766667 -4.8200000 -2.3750000 -1.6100000 -4.5500000
## [25] -3.4350000 -4.7300000 -2.0700000 -2.8200000 -3.0600000 -4.9450000
## [31] -2.8500000 -1.1300000 -4.9850000 -3.5300000 -3.2200000  1.3300000
## [37]  0.3500000 -2.8400000

ferrisiDFqpcr$intensity <- "high"
ferrisiDFqpcr$intensity[ferrisiDFqpcr$delta_ct_cewe_MminusE < -2] <- "middle"
ferrisiDFqpcr$intensity[ferrisiDFqpcr$delta_ct_cewe_MminusE < -4] <- "low"

ggplot(ferrisiDFqpcr,
       aes(x = HI, y = BCI)) +
  geom_point(aes(fill = delta_ct_cewe_MminusE), pch = 21, size = 5) +
  facet_grid(~intensity) +
  geom_smooth() +
  theme_classic()

## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```





Discussed with Stuart:

- Test distributions 0 or counts. Test all vs only infected (“intensity”) distribution. We should be able to fit the distribution of infected on all. Zeros are data. Stochastic move.
- Separation of the zero class. balanced design case/control  $\sim 400 \pm 70$  infecteds SNPchip.
- $H_0$ : no differences are observed
- Separate  $<0.5$  and  $>0.5$  to see the species effect
- timing : WHEN (for my thesis?)