

# Informe del análisis estadístico

TFG Alberto Coll Fernández

30/03/2023

## Introducción y estructura de los datos

Todo el análisis estadístico lo he realizado utilizando R 4.2.2 y RStudio 2022.12.0, y empleando únicamente los datos procedentes de las anémonas que permanecieron en cultivo, dejando los salvajes para otro análisis. El conjunto de datos resultante contaba 36 observaciones de 25 variables. La cabeza del conjunto de datos sería la siguiente:

| individuo | playa     | corte     | tiempo | SOD_t    | CAT_t    | GPx_t      | GR_t     |
|-----------|-----------|-----------|--------|----------|----------|------------|----------|
| 16        | Calahonda | dissected | 0      | 65.68765 | NA       | 5.8131641  | 14.99722 |
| 17        | Calahonda | dissected | 0      | 65.20603 | 28.44396 | 31.7601001 | 12.43759 |
| 18        | Calahonda | dissected | 0      | 88.81272 | 40.73860 | 13.2391590 | 21.27713 |
| 19        | Almuñecar | dissected | 0      | 64.27275 | 27.95511 | 4.4683328  | 14.05074 |
| 20        | Almuñecar | dissected | 0      | 72.94011 | 34.94055 | 14.5441621 | 17.30917 |
| 21        | Almuñecar | dissected | 0      | 69.78500 | 31.30252 | 0.3245675  | 13.24920 |

De estas, las variables relevantes para el ensayo son **playa**, **corte** y **tiempo**. La variable **corte** hace referencia al tratamiento principal del experimento, y toma los valores *control* o *dissected*. Por último, la variable **tiempo** responde a los dos muestreos que se realizaron durante el experimento. No recuerdo ahora mismo en qué mes se realizó cada uno, pero tenía intención de utilizar el nombre del mes para las gráficas y resultados.

El objetivo principal del ensayo es comprobar cómo responden las anémonas a la reproducción asexual en condiciones de cultivo, y cómo interacciona este evento con su maduración sexual en cautividad. Para ello, se midieron distintas enzimas involucradas en sistemas antioxidantes del animal (**SOD**, **CAT**, **GPx**, **GR**, **GST**, **DTD** y **G6PDH**), además de su capacidad antioxidante total (**TEAC**), su concentración de **proteína**, y su concentración de malondialdehído (**MDA**). Estas variables se midieron por duplicado, tanto en los tentáculos del animal como en su tronco o pie, por lo que en total hacen 20 variables de interés.

## Metodología y elección del modelo

Para examinar esta pregunta de investigación, he utilizado un modelo lineal con la siguiente estructura:

$$variable \sim corte * tiempo * playa$$

Que quiere decir que cada una de las variables se modeló como una función de corte, tiempo, playa, y la interacción entre estos factores. El término playa aparece porque el análisis exploratorio desveló que en ocasiones los individuos procedentes de distinta playa seguían tendencias diferentes para determinadas variables.

He optado por este modelo en lugar de hacer varios más sencillos para no realizar tantas comparaciones y test estadísticos, que al final incrementan mucho la probabilidad de error de tipo I. Tras ajustar los modelos, para cada variable, he comprobado si existía un efecto significativo de la interacción corte-tiempo, y, en caso de no haberla, opté por un modelo más parsimonioso, sin ese término de interacción.

Una vez ajustados los modelos, he comprobado el ajuste y las asunciones generales de modelos lineales. La asunción de normalidad de los residuos la he comprobado por medio del test de Shapiro-Wilk, y sólo una variable (GR tentacular) violó esta asunción. Para resolverlo, apliqué una transformación logarítmica. La asunción de homocedasticidad la comprobé visualmente mediante la examinación de la gráfica de Escala-Localización de cada modelo, y no encontré grandes violaciones de esta asunción (hay que tener en cuenta que con tan poco tamaño de muestra siempre hay irregularidades).

Cuando para una variable se detectaron efectos significativos de cualquier tipo, hice un test post-hoc (Test de Tukey) para examinar en detalle dichas diferencias. Por el momento he usado las letras fruto del ANOVA con todos los grupos, pero en algunos casos no se detectan diferencias, así que tengo intención de comparar grupo a grupo como ya hicimos. Ahora mismo aún no aparecen en las gráficas

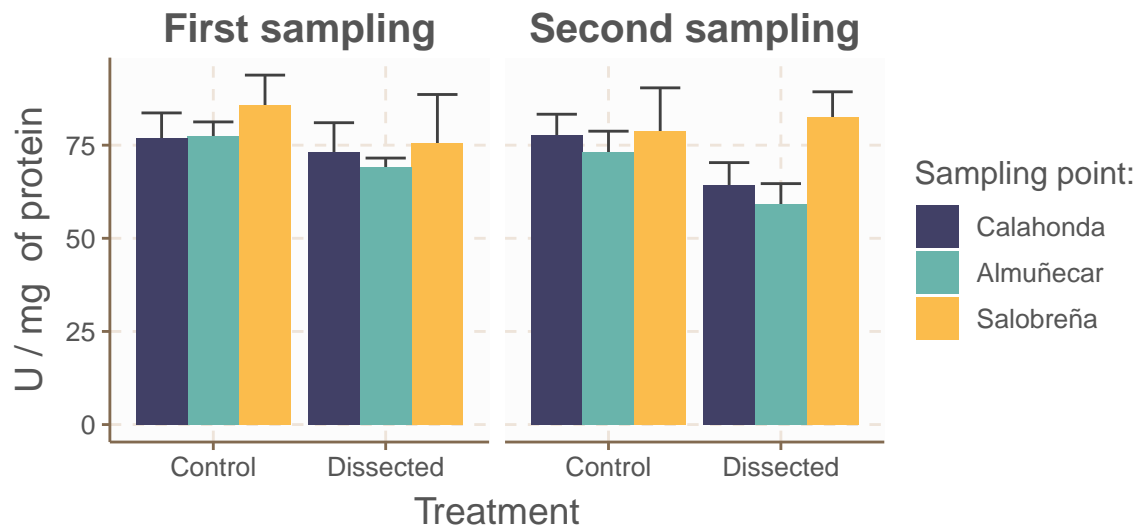
He elaborado dos tipos de gráficas para visualizar los datos. Por un lado, gráficas de barras más tradicionales, y por otro lado, gráficas de interacción, que favorecen la visualización de los efectos de **corte y tiempo**. Los datos utilizados en ambos casos corresponden a la media  $\pm$  error estándar de la media (SEM). Las letras y símbolos de las diferencias significativas aún no están puestos, pero tendré que usar algo más que letras y asteriscos porque tenemos 3 factores, no dos. Pensé en un +, pero coméntame si se suele usar otra cosa porque no lo sé.

## Resultados

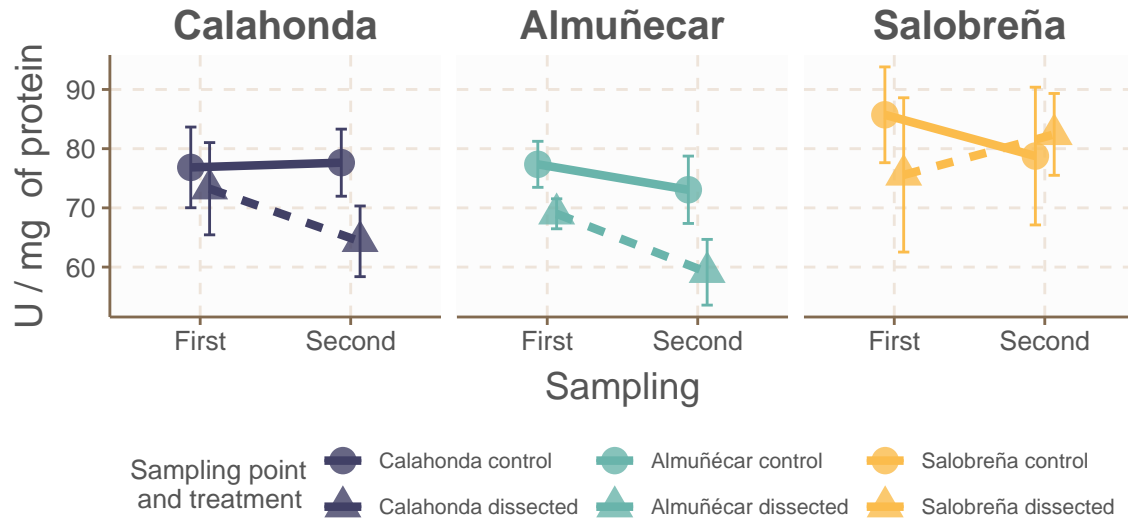
### Superóxido dismutasa (SOD)

#### SOD en tentáculo

En las gráficas de barras, he separado en dos paneles diferentes los datos del primer y segundo muestreo (variable **tiempo**), mientras que en el eje x se encuentra la variable **corte**. Las barras se muestran mediante el código de color. Esta disposición se puede cambiar, así que si quieres alguna modificación en las gráficas basta con decirme qué variable quieres en cada sitio.



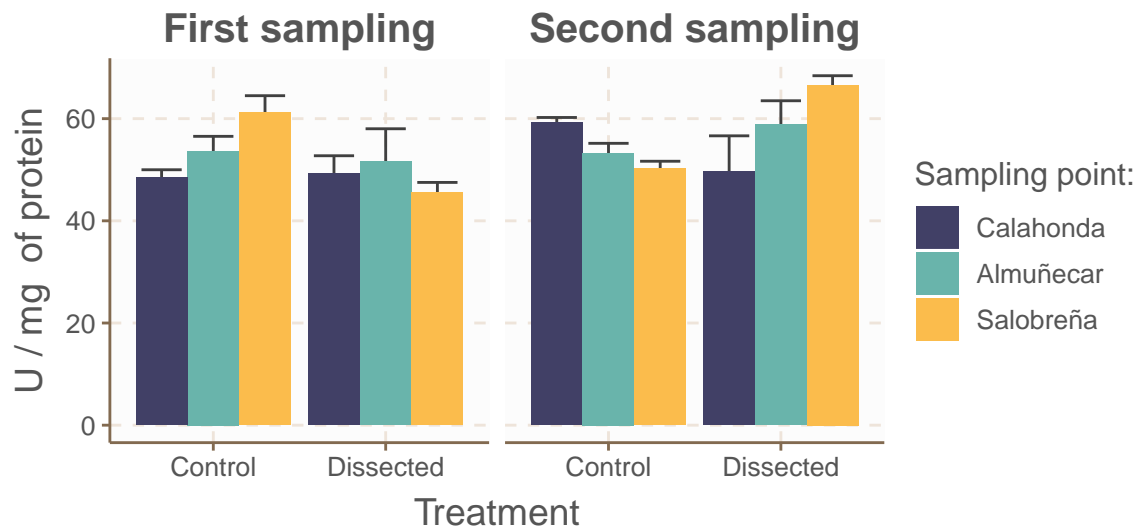
La actividad SOD tentacular se mantuvo relativamente uniforme en todos los grupos experimentales, con una ligera tendencia a mayor actividad en individuos procedentes de Salobreña, pero sin llegar a ser significativas. La gráfica de interacción sería la siguiente:



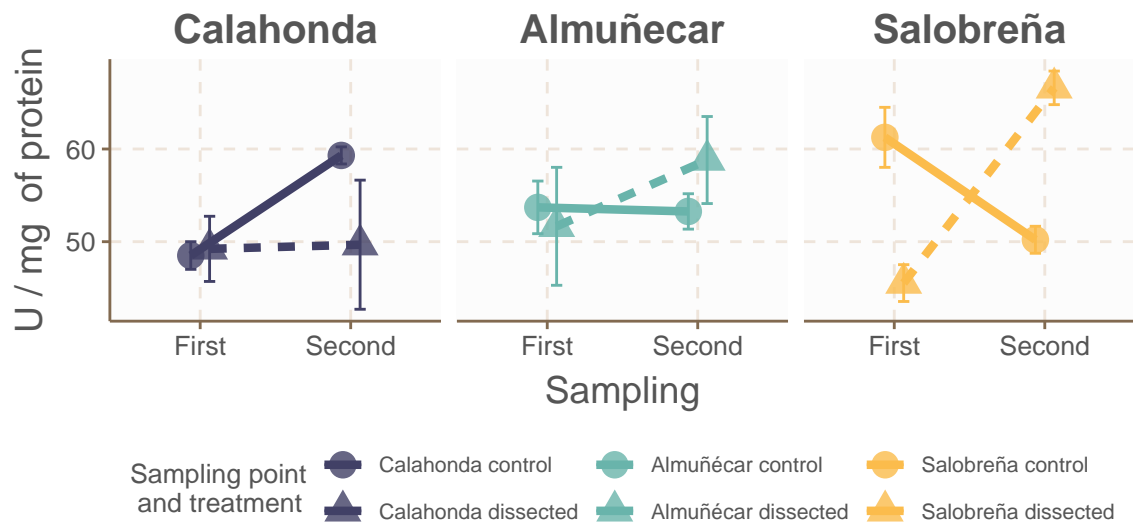
En este tipo de gráficas he separado en tres paneles los datos de cada playa, utilizando además un color distinto para cada una de ellas. Dentro de cada panel, en el eje x tenemos la variable **tiempo**, mientras que el corte lo vemos con dos formas diferentes de cada punto. En esta gráfica podemos ver de nuevo esa tendencia ligeramente diferente en Salobreña.

### SOD en disco pedio

En el caso de la SOD pedia, sí que se detectaron diferencias significativas, en concreto asociadas a la variable **tiempo** y a su interacción con **corte**.



La gráfica de barras para mi gusto es muy liosa, y pienso que se ve bastante mejor la situación en la gráfica de interacción:

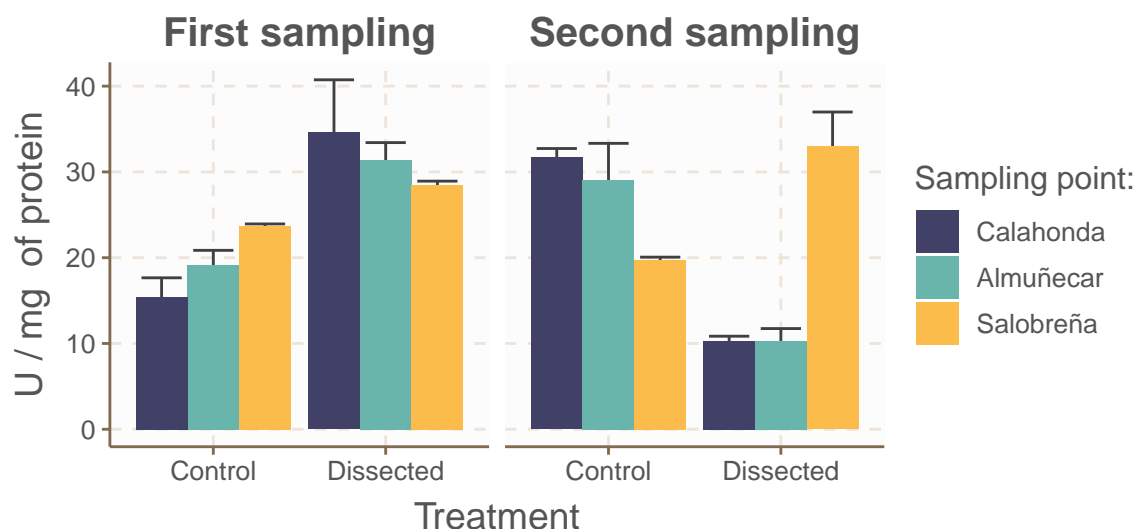


Ahora sí, se puede observar cómo la maduración parece incrementar los niveles de actividad SOD en el disco pedio, sobre todo para los individuos que además fueron cortados, de ahí la significación estadística de la interacción entre ambas variables. Hay que destacar que en Calahonda tenemos el caso contrario, lo que indica que organismos de origen diferente responden de diferente manera a los tratamientos. Aun así, hay que destacar que estas diferencias, aunque significativas, son de pequeña magnitud.

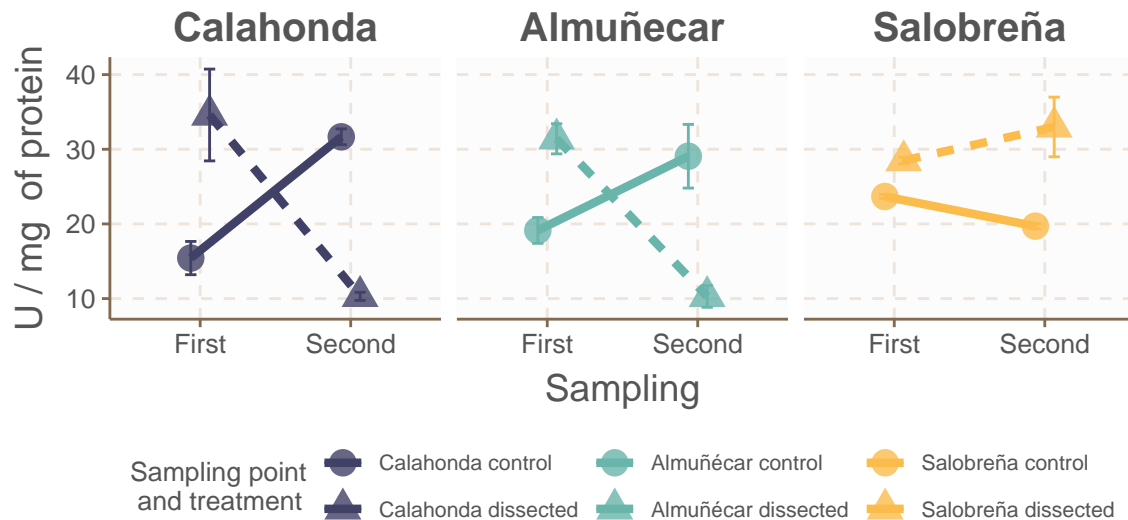
## Catalasa (CAT)

### CAT en tentáculo

En esta enzima sí que encontramos diferencias significativas de mayor relevancia. Vamos primero con la gráfica de barras de la catalasa tentacular:



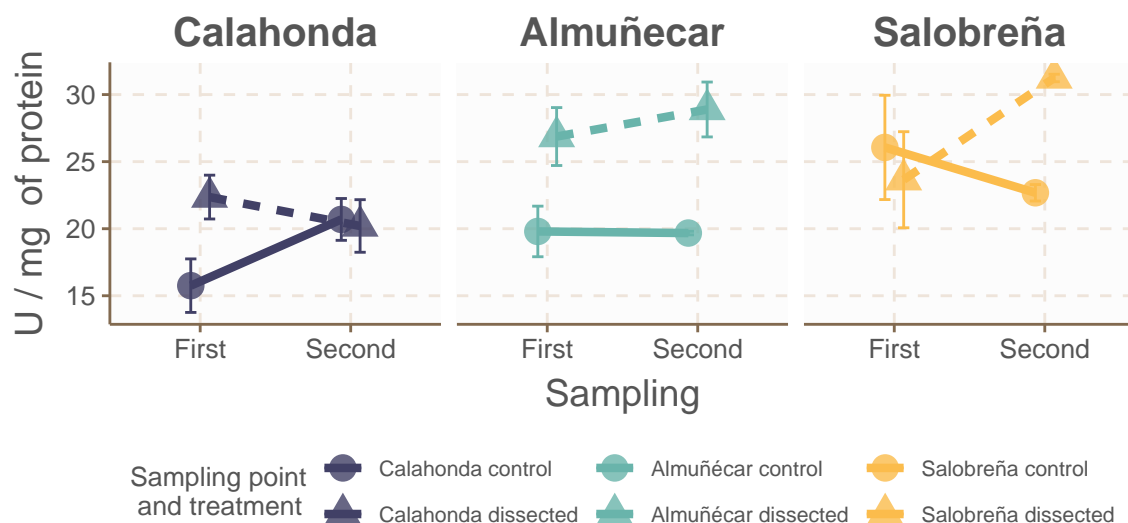
La catalasa tentacular muestra una clara tendencia diferente en los individuos de Salobreña, que se hace evidente en el grupo que fue cortado y maduró sexualmente. A parte de esto, tenemos un efecto de interacción **corte:tiempo**, que se aprecia muy bien en la gráfica de interacción:



Las líneas que se cruzan en Calahonda y Almuñécar indican que, según el individuo haya sido cortados o no, la actividad catalasa responde incrementándose o disminuyendo con la maduración sexual. También se intuye una diferencia en los individuos cortados inmaduros, que fue marginalmente significativa en el modelo estadístico.

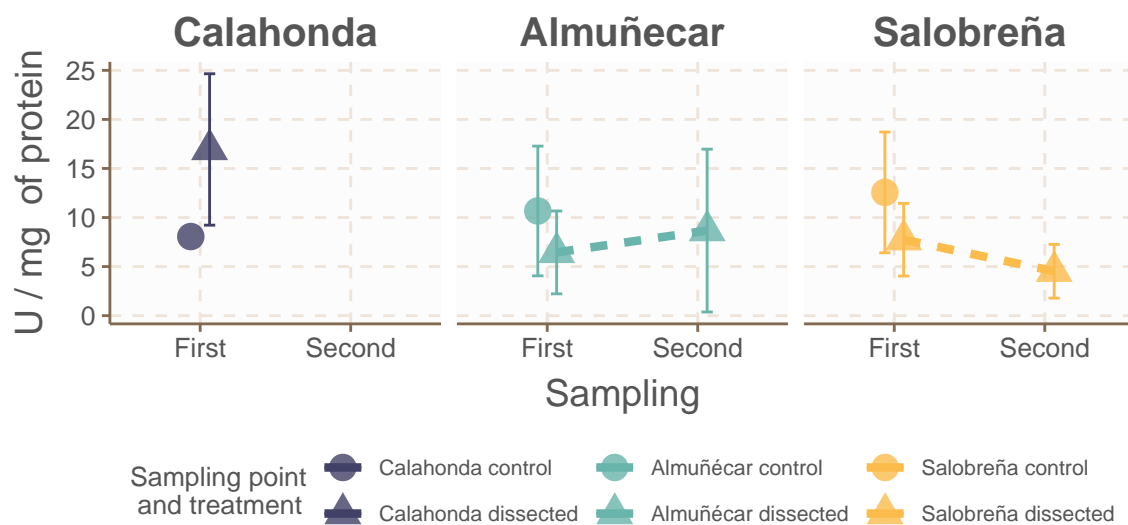
### CAT en disco pedio

En cuanto a la CAT en disco pedio, vemos que el corte produce un efecto muy patente en su actividad, y que también hay diferencias entre las tres playas. En este caso no hubo ningún efecto de la maduración sexual ni ninguna interacción. A partir de ahora voy a omitir la gráfica de barras para no alargar demasiado el documento, y porque creo que es mucho más visual en la de interacción.



## Glutación peroxidasa (GPx)

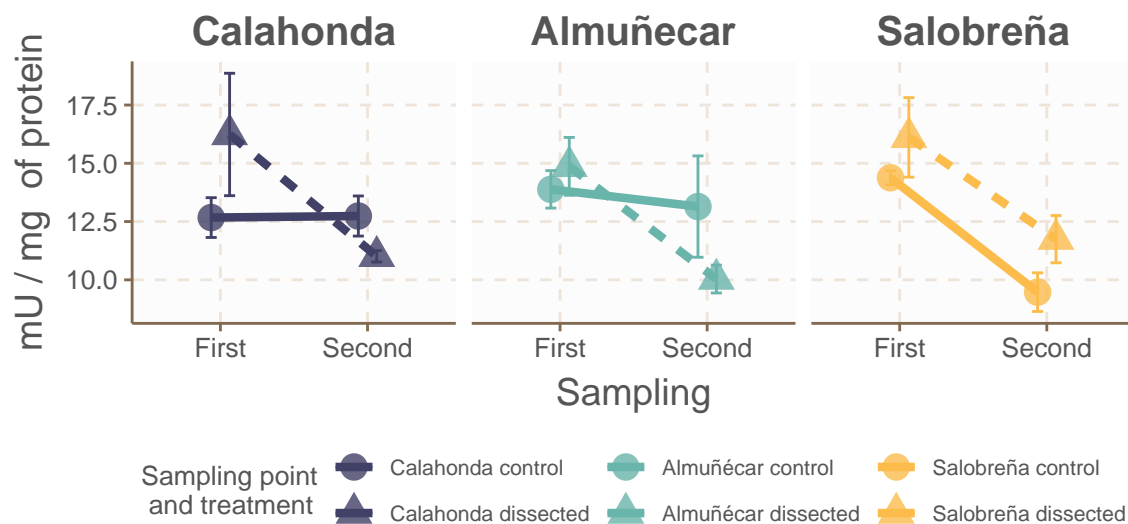
Como ya esperábamos, la GPx no da mucho juego para el análisis porque faltan datos, aquellos en los que obtuvimos valores de actividad negativa. Podría transformarlos en 0 pero no se cómo de fiable sería el análisis. En cualquier caso, faltan datos relativos a grupos experimentales completos, por lo que no he podido ajustar el modelo estadístico. Te dejo aquí una de las gráficas para que veas la situación:



## Glutación reductasa (GR)

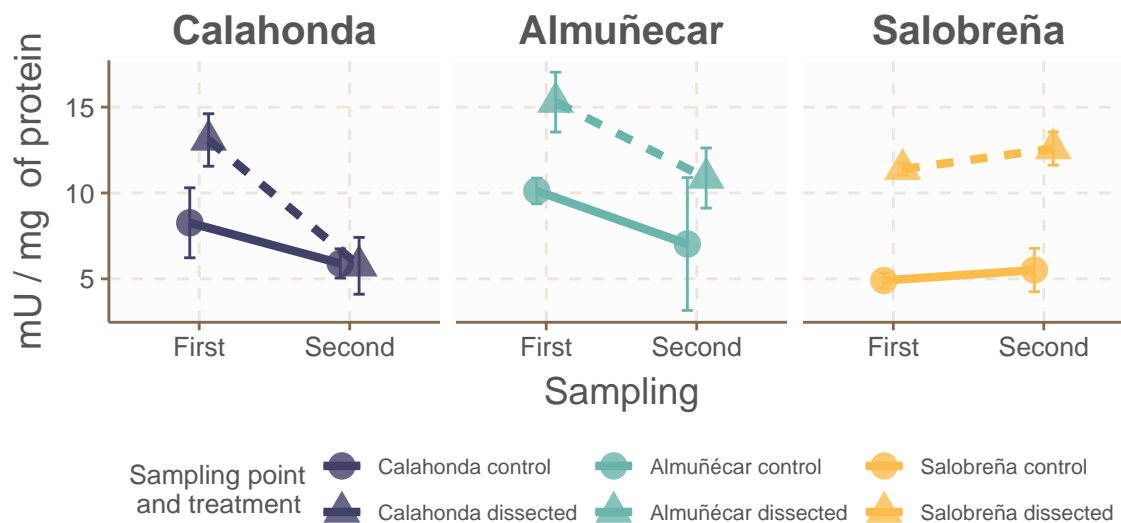
### GR en tentáculo

La actividad GR en tentáculo muestra una respuesta muy clara y uniforme en las tres playas, sin interacciones con otras variables: la actividad disminuye con la maduración sexual:



## GR en disco pedio

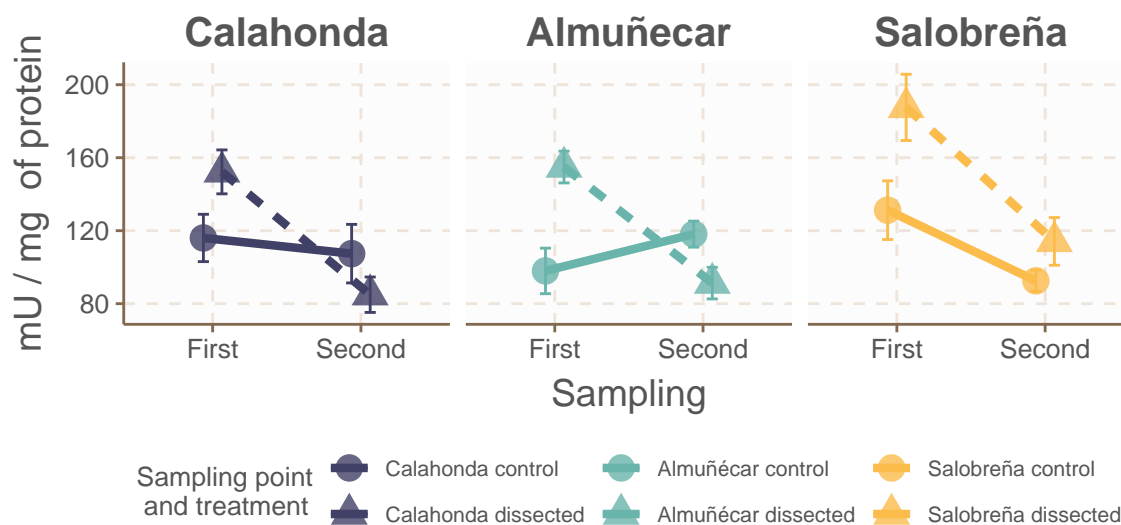
En el caso de su actividad en disco pedio, tenemos además un efecto del corte. De este modo, los individuos cortados tienen siempre mayor actividad GR, y esta descende en todos los casos con la maduración sexual. Si te fijas, el hecho de que las líneas sean paralelas o casi paralelas deja muy claro la ausencia de interacción entre variables.



## Glutathión-S transferasa (GST)

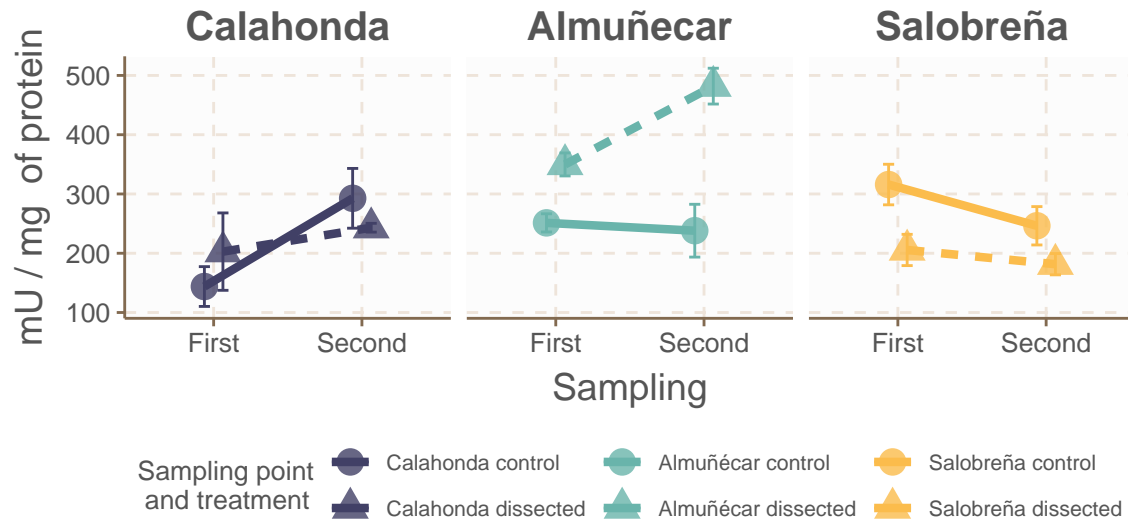
### GST en tentáculo

En el caso de la GST tentacular, su actividad se ve afectada significativamente por el corte y por el tiempo, además de existir una interacción entre ambas. En la gráfica podemos observar que los individuos cortados presentan mayor actividad, pero que esta descende mucho tras la maduración sexual, más que en los organismos que no han recibido dicho tratamiento.



## GST en pie

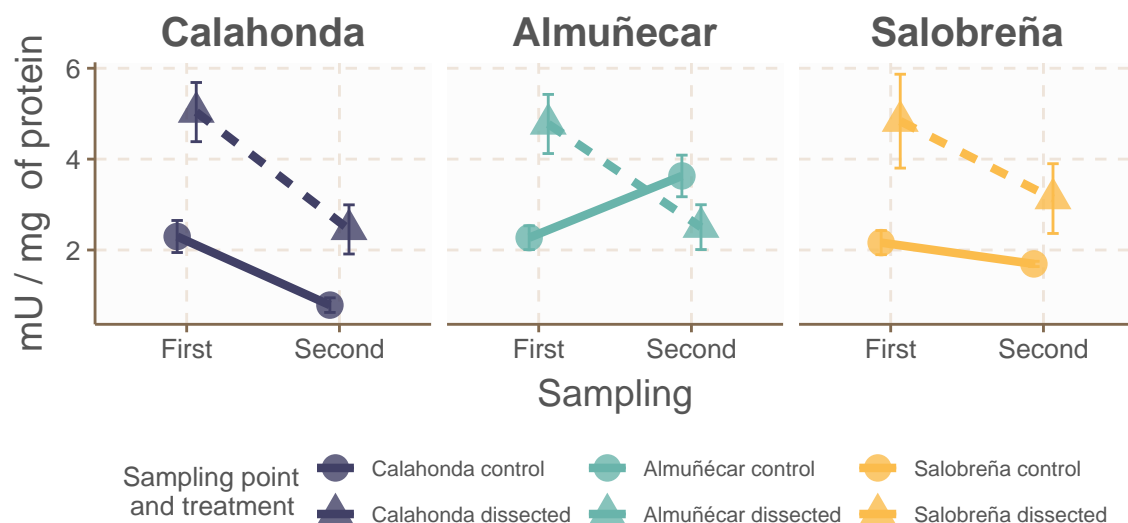
Sin embargo, en la actividad pedia tenemos principalmente diferencias entre puntos de muestreo, y una interacción de la población de origen con la respuesta al corte que vemos muy claramente en Almuñécar, donde la actividad es mucho más alta en individuos cortados. Esta actividad además se incrementa aún más con la maduración sexual, aunque sólo para individuos de este origen.



## DT diaforasa (DTD)

### DTD en tentáculo

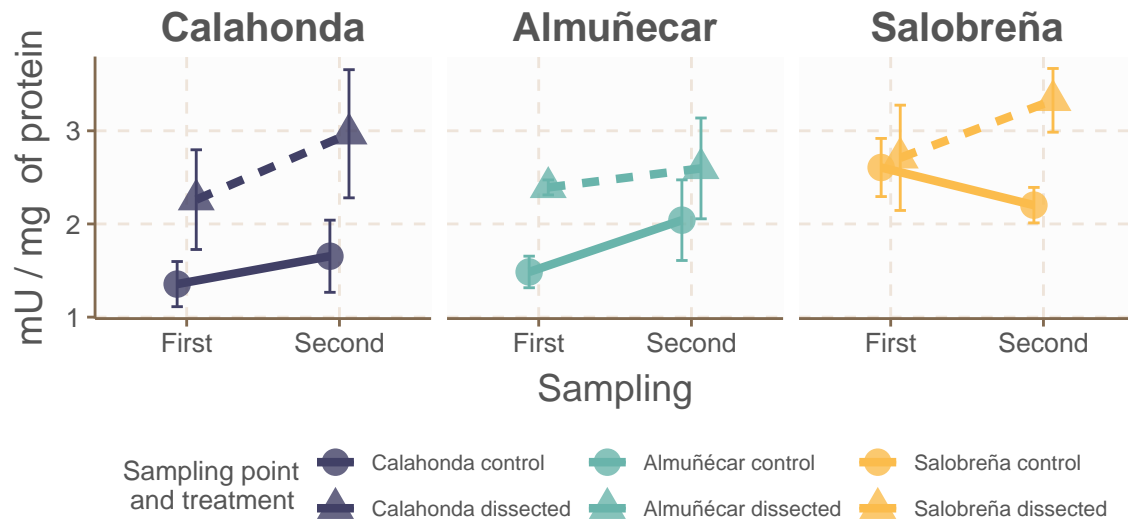
En esta enzima medida en tentáculo volvemos a obtener una respuesta muy clara: los individuos cortados presentan mayor actividad DTD, y esta disminuye en todos los casos con la maduración sexual. El ANOVA avisa que la interacción corte:tiempo es significativa, y podría referirse a que este descenso de la actividad es más acusado en individuos cortados y más suave en el grupo control.





## DTD en disco pedio

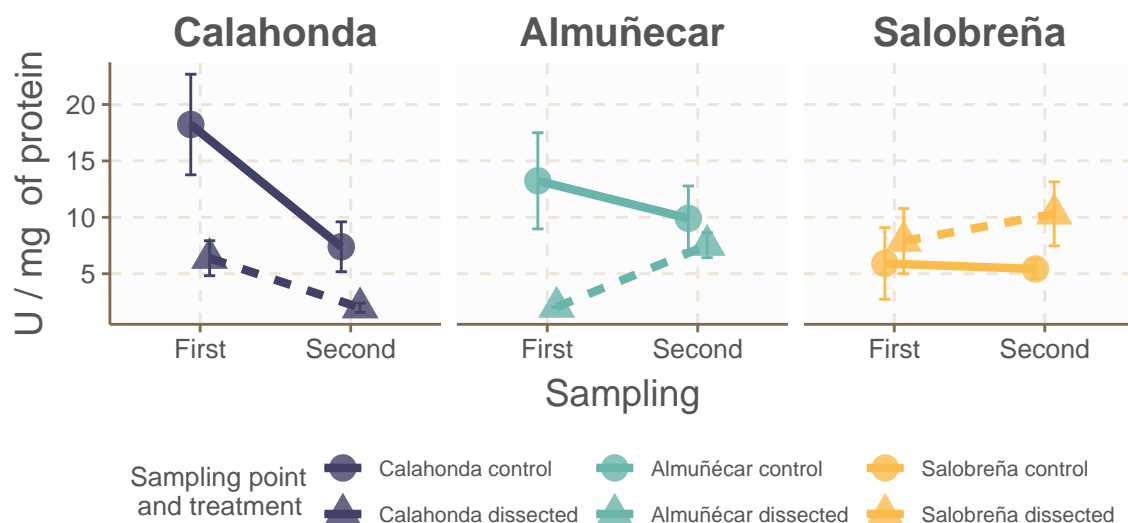
Al medir la misma enzima en el disco pedio, vemos cómo el corte sigue teniendo un efecto importante sobre la actividad DTD, aunque en este caso el efecto de la maduración sexual no es significativo. También parece que Salobreña tiene mayor actividad en general, y el ANOVA ha detectado un efecto significativo de la variable palya (hay que esperar a las letras para ver la naturaleza de estas diferencias).



## Glucosa-6-P deshidrogenasa

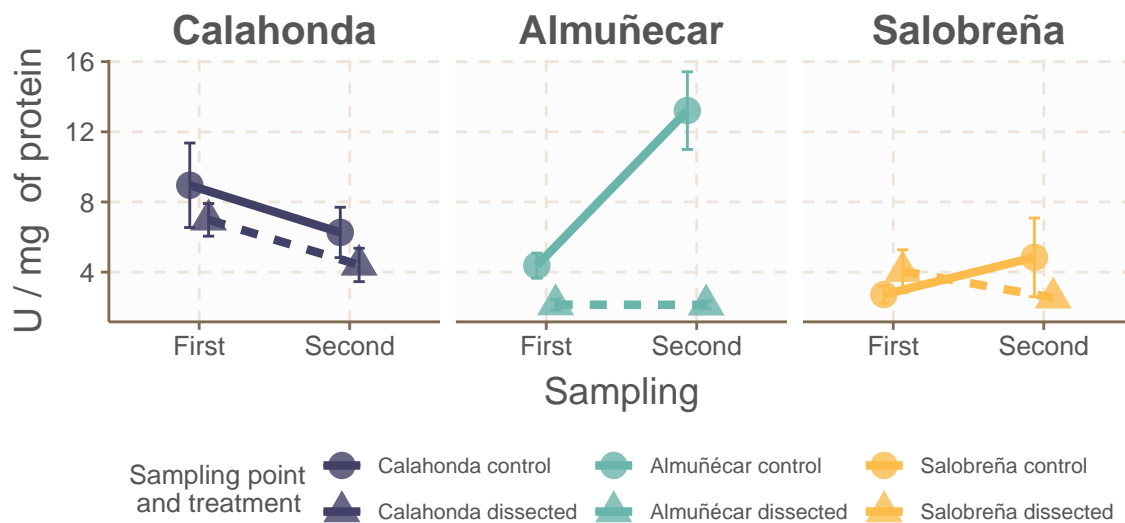
### G6PDH en tentáculo

Con los resultados de esta enzima hay que andarse con ojo porque la calidad de los datos era bastante mejorable, aunque sí que permitió el análisis estadístico. En cualquier caso, la actividad G6PDH tentacular responde reduciéndose en los individuos cortados (a excepción de Salobreña donde parece que no hay mucha diferencia). La maduración sexual no produce diferencias significativas.



## G6PDH en disco pedio

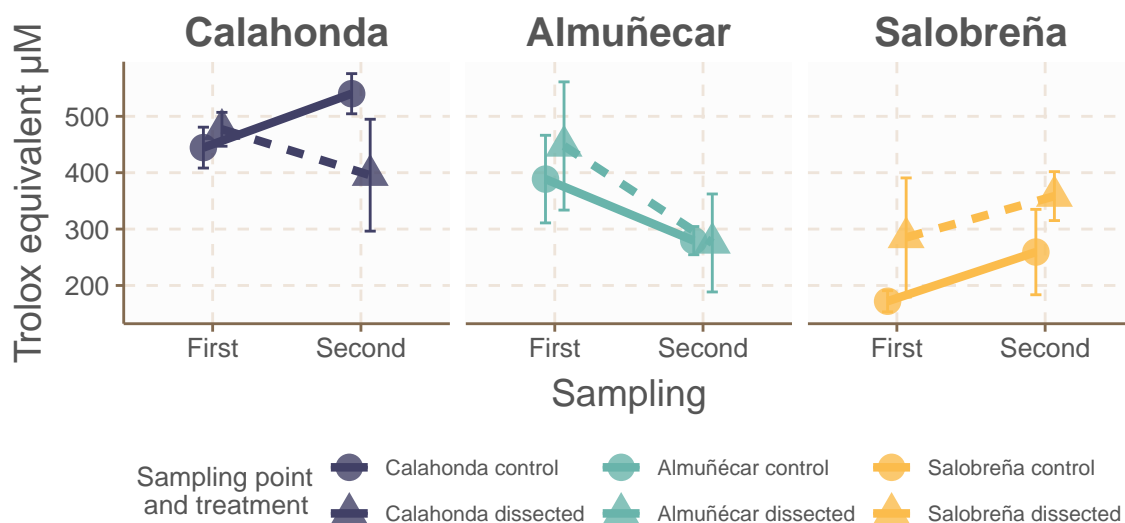
En el caso del pie, nos encontramos con respuestas diferentes según el punto de muestreo. En Calahonda y Salobreña no vemos apenas efectos, y no hay diferencias significativas. Sin embargo, en Almuñécar, la maduración incrementó mucho la actividad G6PDH de los individuos control, mientras que los cortados se mantuvieron estables a través del tiempo.



## Capacidad antioxidante total (TEAC)

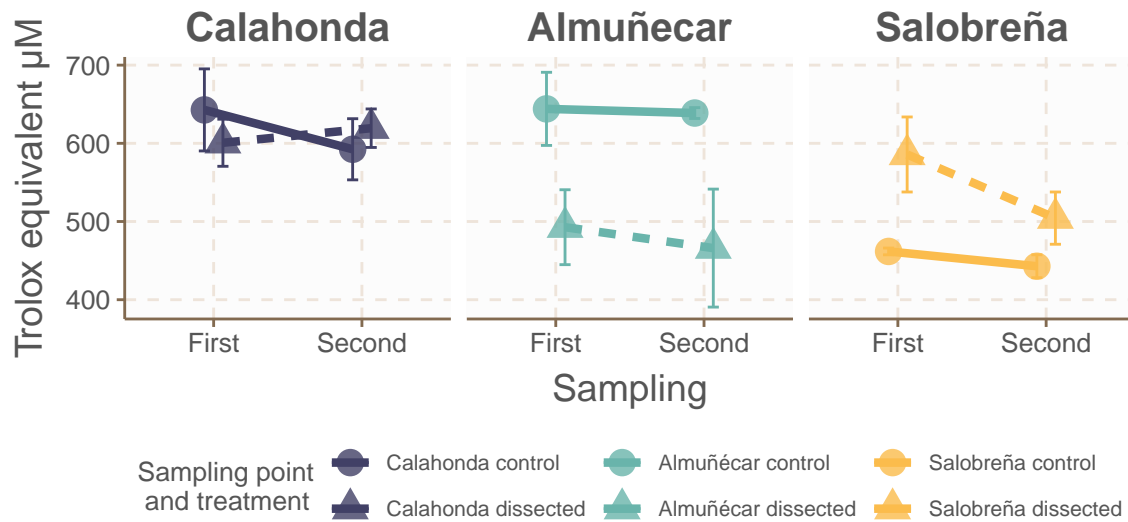
### TEAC en tentáculo

La capacidad antioxidante total en tentáculo no se ve afectado por ninguno de los tratamientos, pero sí que hay diferencias significativas entre playas, donde Salobreña tiene menos capacidad (esperar al test de Tukey para confirmar).



## TEAC en disco pedio

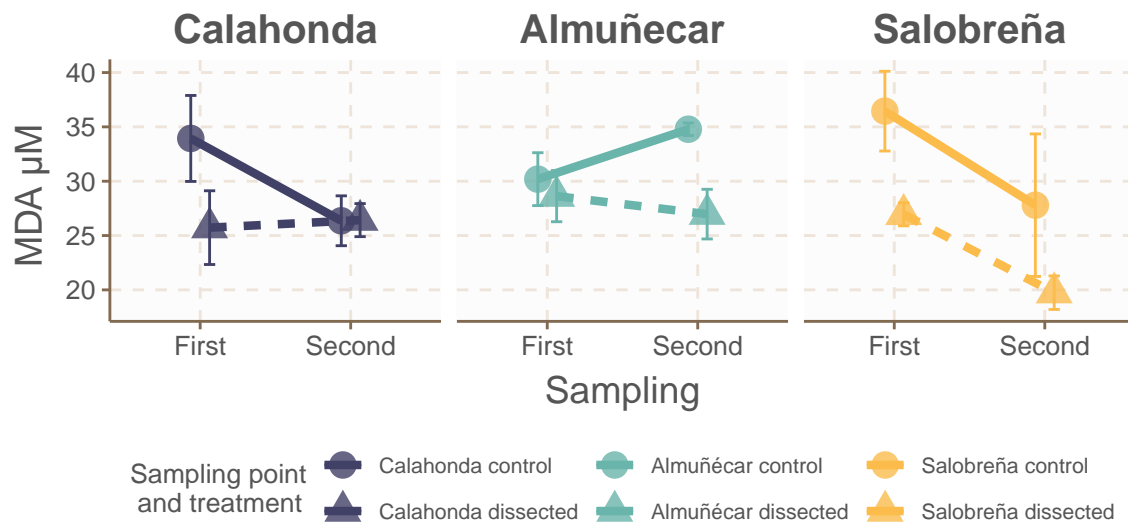
En el pie, volvemos a tener diferencias entre playas, con menores valores en Salobreña y mayores en Calahonda, pero esta vez además vemos un efecto del corte presente sólo en Almuñécar, donde los individuos cortados tienen menor capacidad.



## Malondialdehído (MDA)

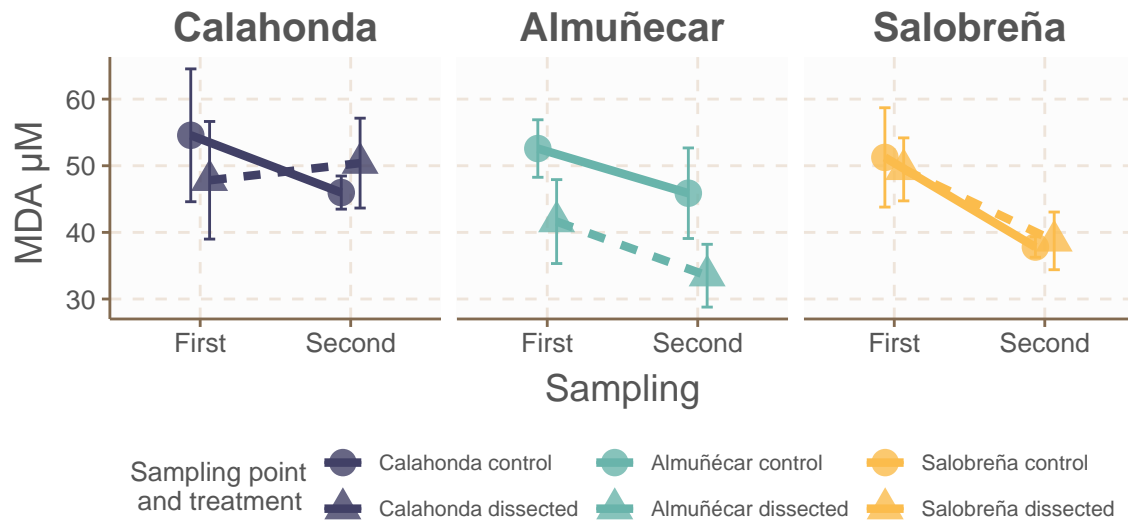
### MDA en tentáculo

La concentración de MDA en tentáculo es bastante uniforme, y aunque el ANOVA destaca un efecto significativo del corte, en la gráfica prácticamente no se observa. Es curioso que los individuos control, sin cortar, tienen mayor concentración de MDA.



## MDA en disco pedio

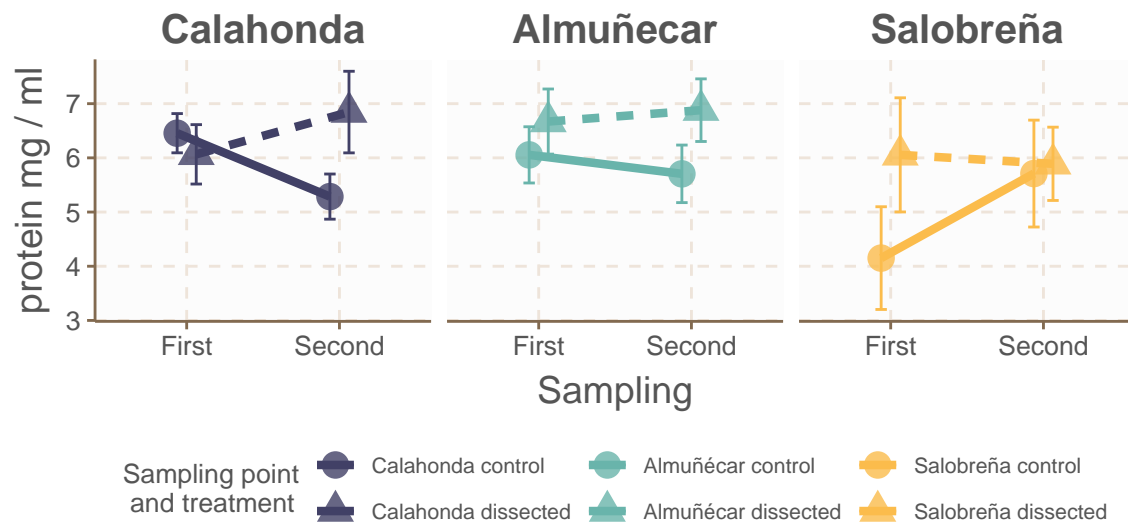
En el caso del pie, volvemos a tener una gran uniformidad en concentración de MDA, y en este caso el ANOVA destaca un efecto algo significativo de la maduración, que parece hacer descender la concentración de este compuesto en todos los casos.



## Proteína

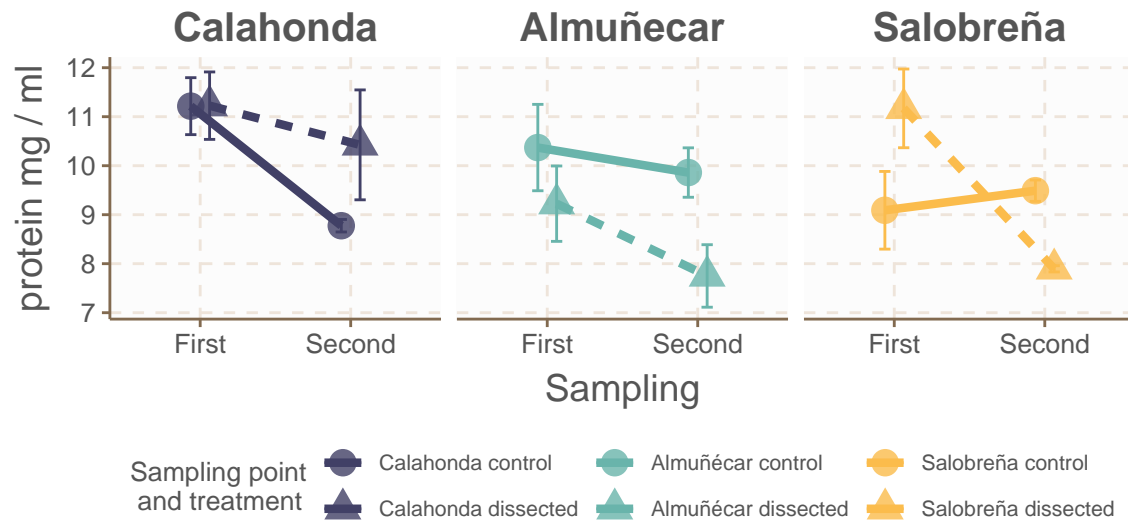
### Proteína en tentáculo

La proteína no es de mucho interés salvo para expresar las unidades de actividad enzimática, pero tenemos su análisis igual. Vemos que prácticamente no hay diferencias, aunque el ANOVA destaca un pequeño efecto significativo del corte.



## Proteína en disco pedio

En el caso del pie, la única tendencia estadísticamente significativa que se observa es que la concentración de proteína tiene a disminuir ligeramente con la maduración, pero en las gráficas vemos que es algo sutil.



## Conclusiones

- La enzima GPx no fue ni siquiera analizable, mientras que la G6PDH sí que permitió el análisis pero con un ajuste muy malo, por lo que sus conclusiones podrían no ser las más fiables.
- Las variables que más bondad de ajuste ( $R^2_{ajustado}$ ) presentan son la CAT tentacular, ambas GR, ambas GST y ambas DTD. Esto quiere decir que para esas enzimas, los factores playa, corte y tiempo consiguen explicar una buena parte de la variación que presentan, y por tanto el modelo es muy bueno “prediciendo” sus valores.
- Es también en estas variables donde encontramos más cambios asociados al corte y a la maduración, y donde más claramente se observan. En estas enzimas, los individuos cortados presentaban mayores niveles de actividad, y en algunas de ellas estos niveles descendían bruscamente con la maduración sexual.
- La SOD, MDA y TEAC presentaron en general muy poca variación a lo largo de los tratamientos, con sólo unos pocos efectos significativos. En general no son parámetros que se alteren mucho al inducir la reproducción asexual o incluso con la maduración sexual de las anémonas.