Вообще в начале должна быть модель 5 (предсказания Ptros по PT). Потом сказать, что мы пишем уравнения по коэффициентам для модели 5 и модели 3 и рекомендуем их использовать в Белом море и опресненной части Кольского залива. А потом уже вот это, для других пользователей:

1. Если у вас уже есть большой массив генотипированных данных из гибридных зон МЕ и МТ (в максимально широком диапазоне варьирования Ptros) (**давайте МТ и МГалло тоже предлагать попробовать**, ссылаясь на литературные данные) и их раковины, вы можете сделать себе такой же анализ, как у нас: подобрать регрессионные модели, связывающие частоты одного из генотипов (зависимая переменная) с частотой характерного морфотипа (предиктор) и использовать их, чтобы в дальнейшем не генотипировать.
2. Вы можете полностью повторить наш анализ: описать взаимосвязь частот генотипов с частотами морфотипов на фоне максимально возможного варьирования таксономической структуры выборок.
3. Если у вас нет данных для пункта 1 или вам не хочется тратить время и деньги на пункт 2, но хочется быстро и минимально затратно прикинуть, как по морфотипам оценить генетическую структуру и вероятность корректного определения особей до вида по морфотипам, то логика анализа следующая:

- Соберите калибровочные выборки для оценки следующих вероятностей:

**P(T|MT)c** – вероятность встретить Т-морфотип у МТ

**P(T|ME)c** – вероятность встретить Т-морфотип у МE

Ниже мы напишем, как мы выясняли, каких выборок и в каком количестве минимально достаточно брать.

- Генотипируйте их и определите морфотипы.

- Опишите зависимости, которые выводятся согласно теореме Байеса (это и есть калькуляторы).

Калькулятор 1. Первая зависимость отражает связь **Ptros** с частотой T-морфотипа в изучаемой выборке (**РТ**) и выводится из формулы полной вероятности, отражающей вероятность появления события «Т-морфотип»:

**PT =** **Ptros\*P(T|MT)c + (1 - Ptros)\*P(T|MЕ)c**

Отсюда получаем следующую зависимость:

**=**

Приведенная формула накладывает ряд ограничений:

1. Если в калибровачных выборках вы получили **P(T|ME)с** = **P(T|MT)с** или **P(T|ME)с** > **P(T|MT)с**, то применение этой формулы невозможно, тк в этих случаях T-морфотип не является диагностичным маркером для МТ или вообще является маркером для МЕ, соответственно.
2. Использование **P(T|ME)с** и **P(T|MT)с,** полученных по калибровочным выборкам, предполагает, что в регионе, в котором проходит исследование, наблюдаемые значения частот Т-морфотипа в выборках (**РТ**) находятся в пределах от **P(T|MЕ)с** до **P(T|MT)с**включительно.
3. Если при дальнейшем использовании формулы будут обнаружены выборки с иной **РТ** (см. п.2) нужно откорректировать калькулятор по новым данным (отгенотипировать эту выборку и пересчитать калибровочные значения)

Калькулятор 2. Следующие зависимости отражают связь вероятности корректного определения особей до вида по морфотипам с **Ptros** и высчитываются по теореме Байеса:

**P(MT|T) =**

**P(ME|E) =**

Значение **Ptros** получаем из калькулятора 1.

**Как подобрать калибровочные выборки?**

1. На нашем массиве данных мы проводим анализ, который даст ответ на вопрос, какой минимально достаточный набор выборок надо брать для калькуляторов.

- ну и тут вот про бутстрепирование максимально кратко написать.