

Methods to resolve this: Convert Y to (C or T)

| | | | | | | | | | | | | |
|-----------------------------|---|---|---|---|---|---|---|---|---|---|---|--|
| lcl consensus | A | A | A | G | A | A | Y | G | T | C | A | |
| gb OQ963804.1 [organism=Ir | A | A | A | G | A | A | C | G | T | C | A | |
| gb OQ958820.1 [organism=Ir | A | A | A | G | A | A | T | G | T | C | A | |
| gb OQ958812.1 [organism=Ir | A | A | A | G | A | A | T | G | T | C | A | |
| gb OQ958804.1 [organism=Ir | A | A | A | G | A | A | T | G | T | C | A | |
| gb OQ961564.1 [organism=Ir | A | A | A | G | A | A | T | G | T | C | A | |
| gb OQ963300.1 [organism=Ir | A | A | A | G | A | A | C | G | T | C | A | |
| gb OQ961548.1 [organism=Ir | A | A | A | G | A | A | C | G | T | C | A | |
| gb OQ961556.1 [organism=Ir | A | A | A | G | A | A | C | G | T | C | A | |

What is the problem with this approach?

| | | | | | | | | | | | | | | | | | | | | | | | |
|-----------------------------|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| lcl consensus | A | A | A | G | A | A | T | G | T | C | A | C | T | G | T | T | A | C | A | C | A | T | G |
| gb OQ963804.1 [organism=Ir | A | A | A | G | A | A | C | G | T | C | A | C | T | G | T | T | A | C | A | C | A | T | G |
| gb OQ958820.1 [organism=Ir | A | A | A | G | A | A | T | G | T | C | A | C | T | G | T | T | A | C | A | C | A | T | G |
| gb OQ958812.1 [organism=Ir | A | A | A | G | A | A | T | G | T | C | A | C | T | G | T | T | A | C | A | C | A | T | G |
| gb OQ958804.1 [organism=Ir | A | A | A | G | A | A | T | G | T | C | A | C | T | G | T | T | A | C | A | C | A | T | G |
| gb OQ961564.1 [organism=Ir | A | A | A | G | A | A | T | G | T | C | A | C | T | G | T | T | A | C | A | C | A | T | G |
| gb OQ963300.1 [organism=Ir | A | A | A | G | A | A | C | G | T | C | A | C | T | G | T | T | A | C | A | C | A | T | G |
| gb OQ961548.1 [organism=Ir | A | A | A | G | A | A | C | G | T | C | A | C | T | G | T | T | A | C | A | C | A | T | G |
| gb OQ961556.1 [organism=Ir | A | A | A | G | A | A | C | G | T | C | A | C | T | G | T | T | A | C | A | C | A | T | G |

Methods to resolve this: Convert Y to (C or T)

| | | | | | | | | |
|-----------------------------|---|---|---|---|---|---|---|---|
| lcl consensus | A | A | A | G | A | A | Y | G |
| gb OQ963804.1 [organism=Ir | A | A | A | G | A | A | C | G |
| gb OQ958820.1 [organism=Ir | A | A | A | G | A | A | T | G |
| gb OQ958812.1 [organism=Ir | A | A | A | G | A | A | T | G |
| gb OQ958804.1 [organism=Ir | A | A | A | G | A | A | T | G |
| gb OQ961564.1 [organism=Ir | A | A | A | G | A | A | T | G |
| gb OQ963300.1 [organism=Ir | A | A | A | G | A | A | C | G |
| gb OQ961548.1 [organism=Ir | A | A | A | G | A | A | C | G |
| gb OQ961556.1 [organism=Ir | A | A | A | G | A | A | C | G |

We don't know the ground truth, this introduces bias into our dataset.

This approach does not scale.

Builds bad habits.

| | | | | | | | | | | | | | | | | | | | | | | | |
|-----------------------------|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| lcl consensus | A | A | A | G | A | A | T | G | T | C | A | C | T | G | T | T | A | C | A | C | A | T | G |
| gb OQ963804.1 [organism=Ir | A | A | A | G | A | A | C | G | T | C | A | C | T | G | T | T | A | C | A | C | A | T | G |
| gb OQ958820.1 [organism=Ir | A | A | A | G | A | A | T | G | T | C | A | C | T | G | T | T | A | C | A | C | A | T | G |
| gb OQ958812.1 [organism=Ir | A | A | A | G | A | A | T | G | T | C | A | C | T | G | T | T | A | C | A | C | A | T | G |
| gb OQ958804.1 [organism=Ir | A | A | A | G | A | A | T | G | T | C | A | C | T | G | T | T | A | C | A | C | A | T | G |
| gb OQ961564.1 [organism=Ir | A | A | A | G | A | A | T | G | T | C | A | C | T | G | T | T | A | C | A | C | A | T | G |
| gb OQ963300.1 [organism=Ir | A | A | A | G | A | A | C | G | T | C | A | C | T | G | T | T | A | C | A | C | A | T | G |
| gb OQ961548.1 [organism=Ir | A | A | A | G | A | A | C | G | T | C | A | C | T | G | T | T | A | C | A | C | A | T | G |
| gb OQ961556.1 [organism=Ir | A | A | A | G | A | A | C | G | T | C | A | C | T | G | T | T | A | C | A | C | A | T | G |