Results report

Elisabetta Canteri

With this report I aim to present the results obtained from the calculation of the genetic diversity of birds, using both Cytochrome-b and CO1 sequences.
We used two different datasets for both loci, one with all the available geo-referenced sequences and one with only the sequences falling inside the breeding range. We use the range of the species to evaluate the accuracy of the coordinates of the sequences; however, compared to mammals and amphibians, birds' mobility creates additional complexity, as their range is partitioned in a breeding range and a wintering range. The breeding range is the one of which we have more knowledge, therefore this is the one that we use to determine the accuracy of the sequences. Nevertheless, if our dataset is composed only by the sequences falling inside the breeding range, we finish loosing a lot of data, possibly affecting the results. For this reason we decided to create these two datasets and to calculate the genetic diversity for both of them.

Hereby are presented the results of the calculation. We calculate the genetic diversity per grid cell and per latitudinal bands.

Cytochrome-b

The first two figures refer to genetic diversity calculations using all the available georeferenced sequences of cytochrome-b. The same was performed using only the sequences inside the breeding range, which result is portrayed in Figure 3 and 4. Finally, the results per latitudinal band of the two datasets are plotted in the same graph in Figure 5.

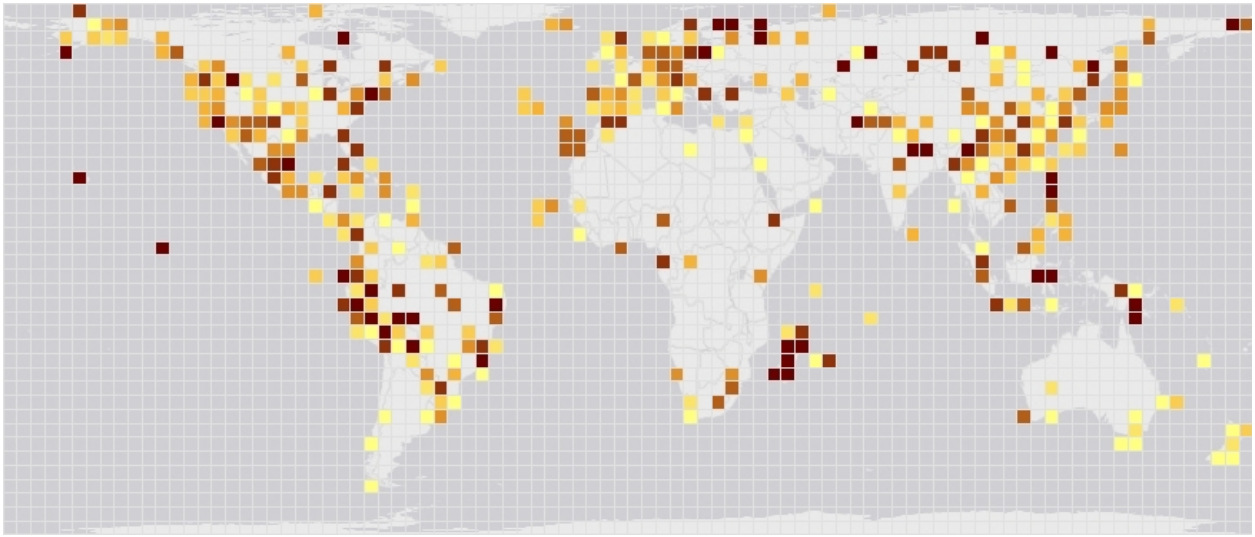

Figure 1: The figure shows the global distribution of genetic diversity using all georeferenced sequences. The map has a grid cell resolution of 4x4 and darker colours mean higher genetic diversity.

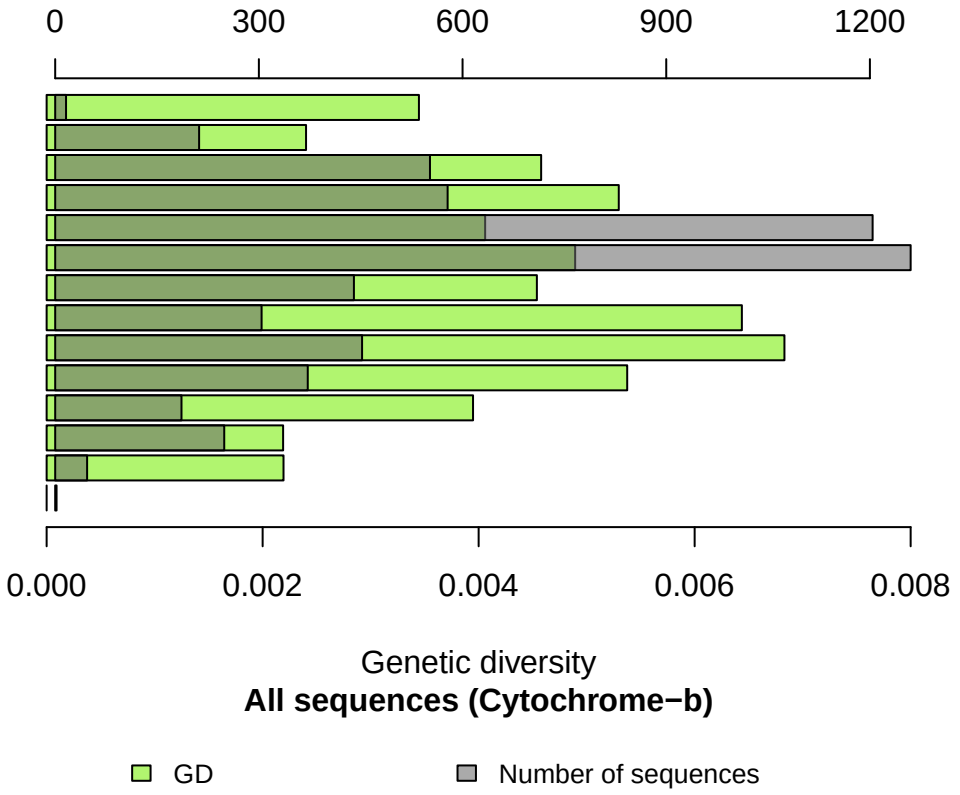

Figure 2: The figure shows the getic diversity (GD) and the number of sequences per latitudinal band. All georeferenced cytochrome-b sequences were used for the calculation of GD.

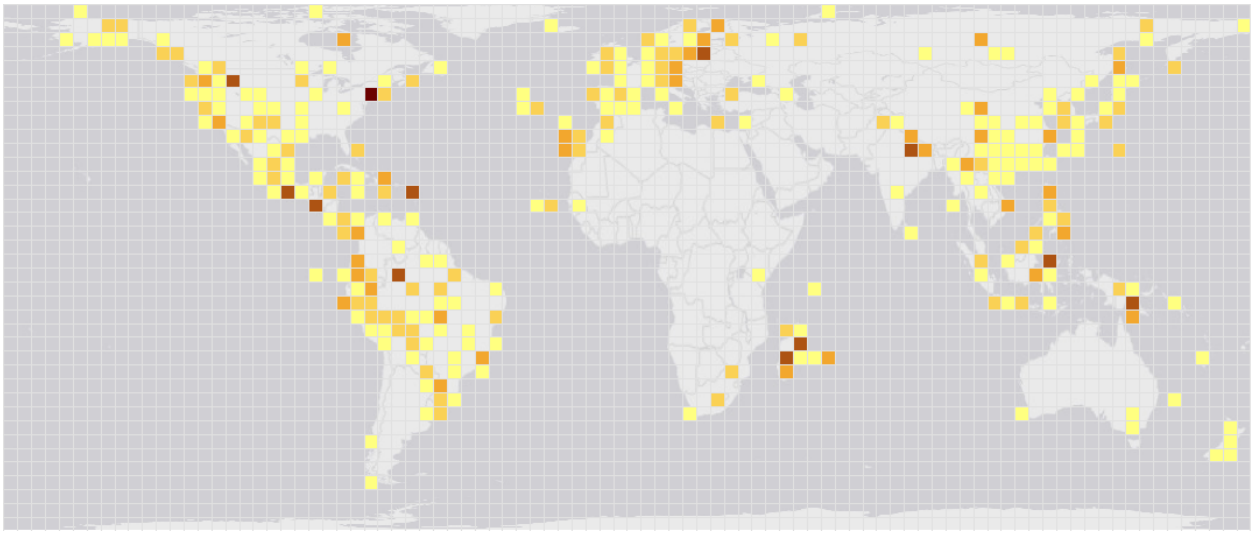

Figure 3: The figure shows the global distribution of genetic diversity using only the georeferenced sequences falling inside the breeding range. The map has a grid cell resolution of 4x4 and darker colours mean higher genetic diversity.

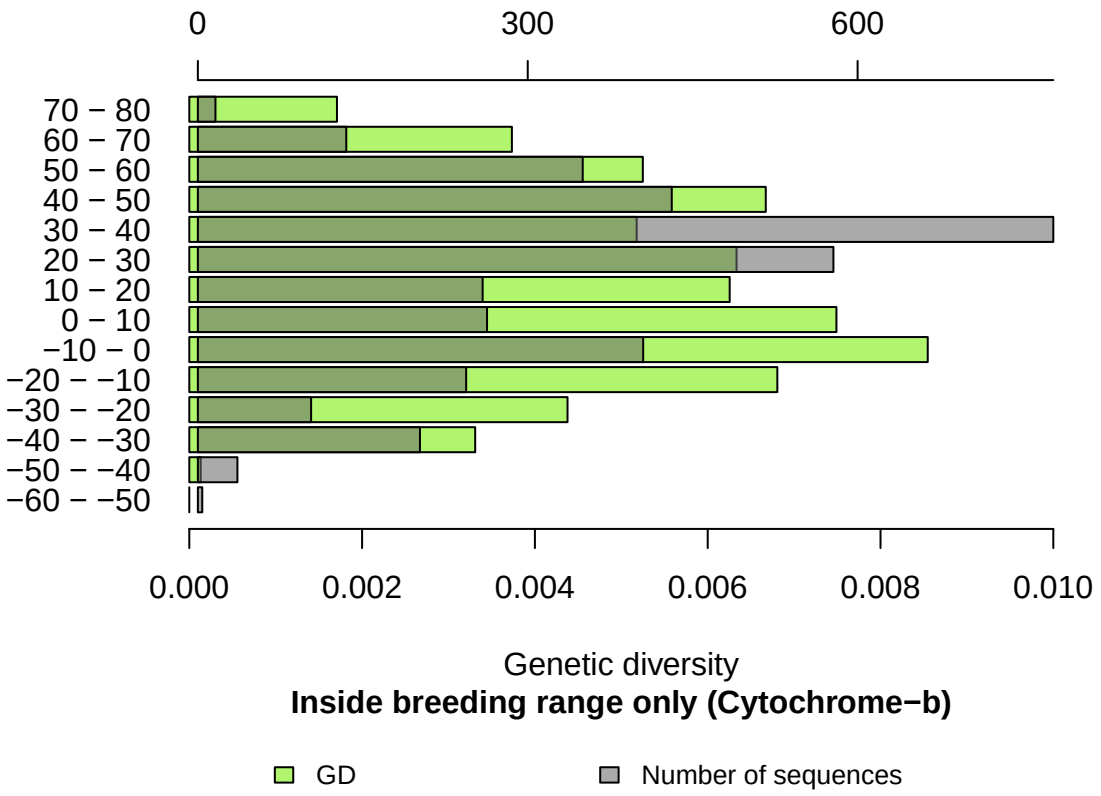

Figure 4: Here is represented the genetic diversity and the number of sequences per latitudinal band. GD was calculated using only the sequences inside the breeding range.

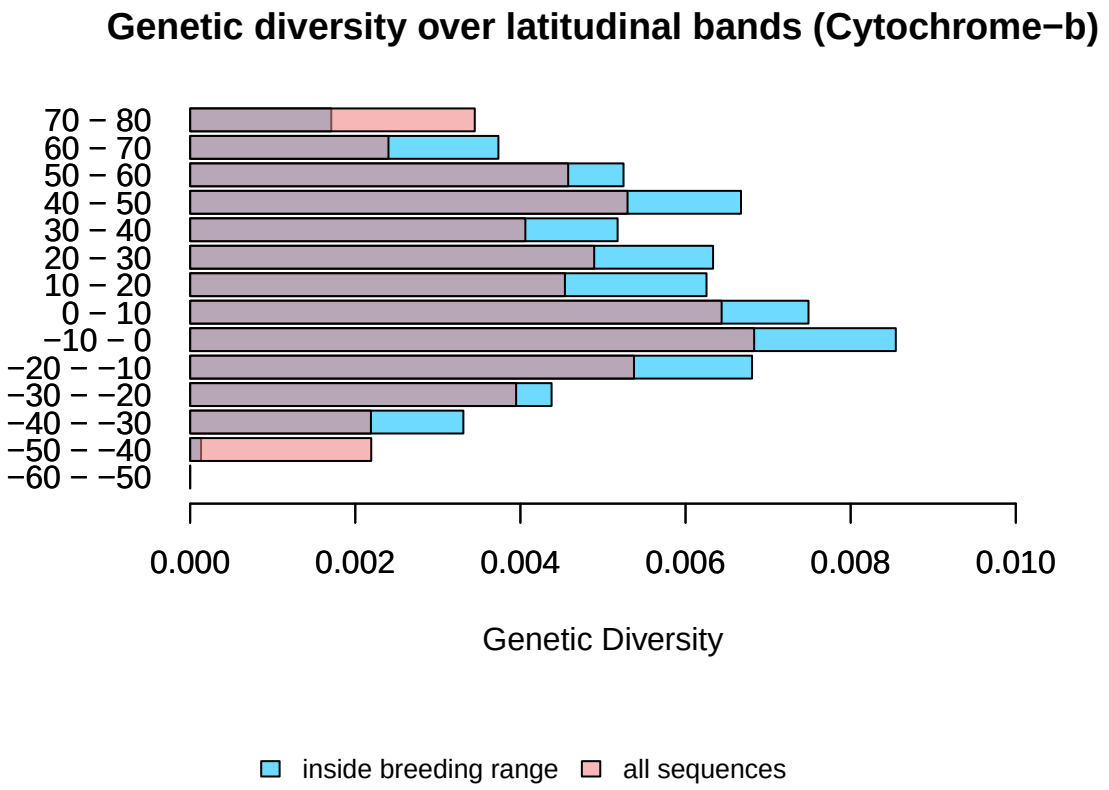

Figure 5: The figure shows the sovrapposition of results per latitudinal band obtained from the two different datasets.

CO1

The same figures are created, and here presented, after performing the same calculations for the CO1 sequences.

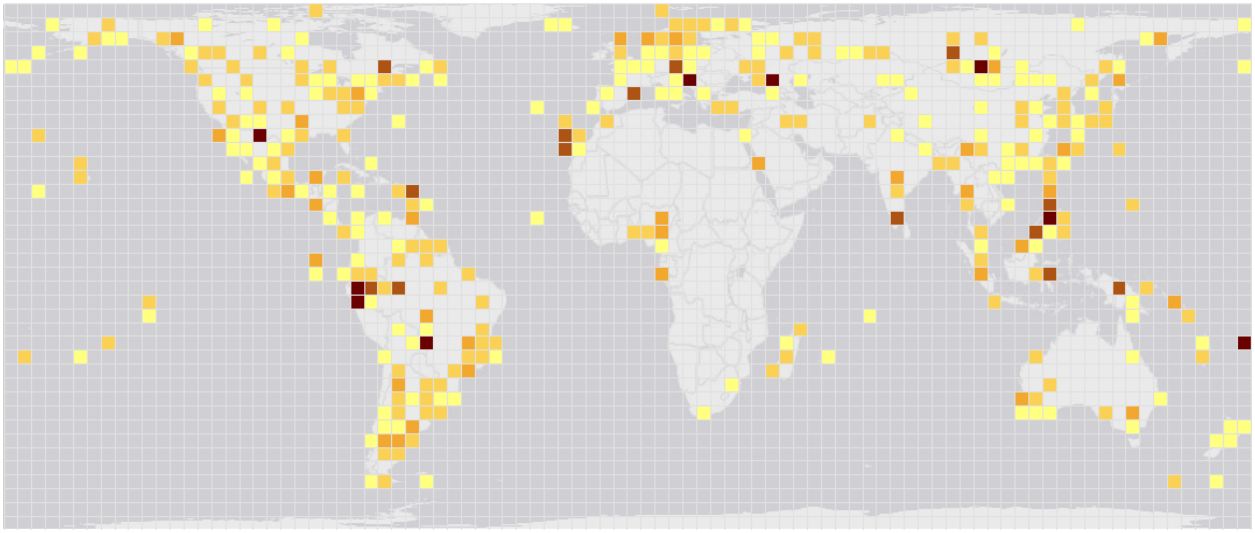

Figure 6: The figure shows the global distribution of genetic diversity using all georeferenced sequences. The map has a grid cell resolution of 4x4 and darker colours mean higher genetic diversity.

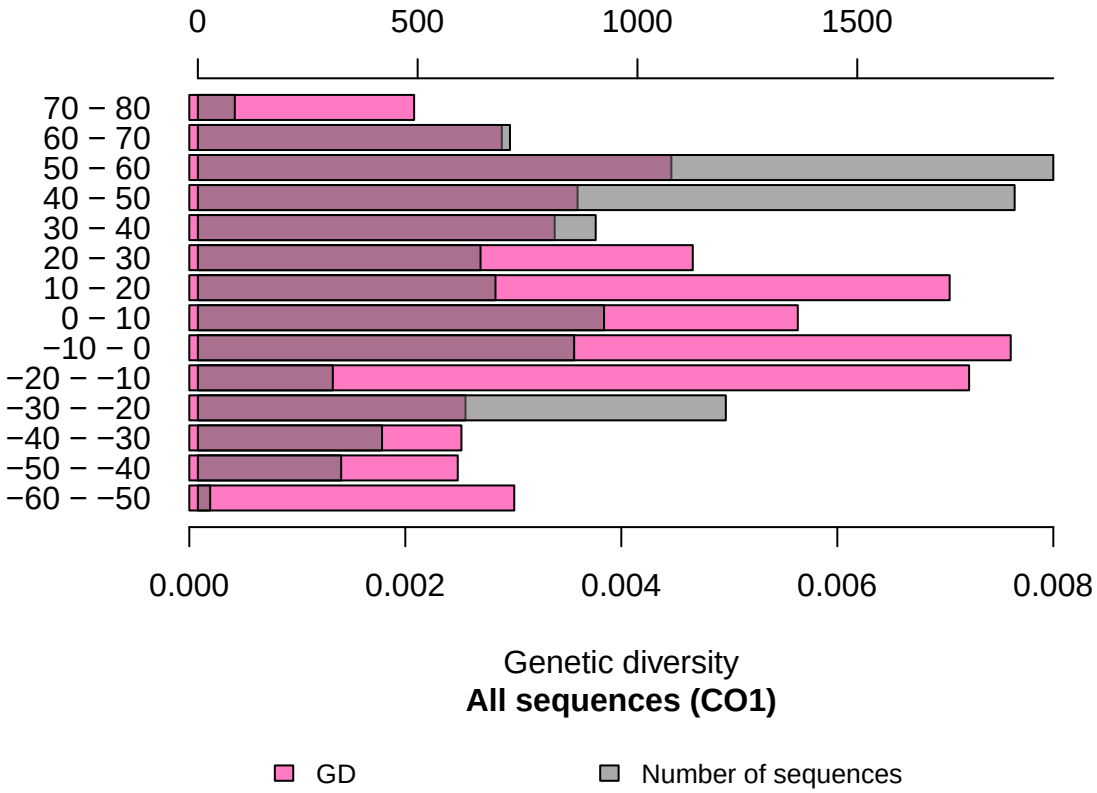

Figure 7: The figure shows the getic diversity (GD) and the number of sequences per latitudinal band. All georeferenced CO1 sequences were used for the calculation of GD.

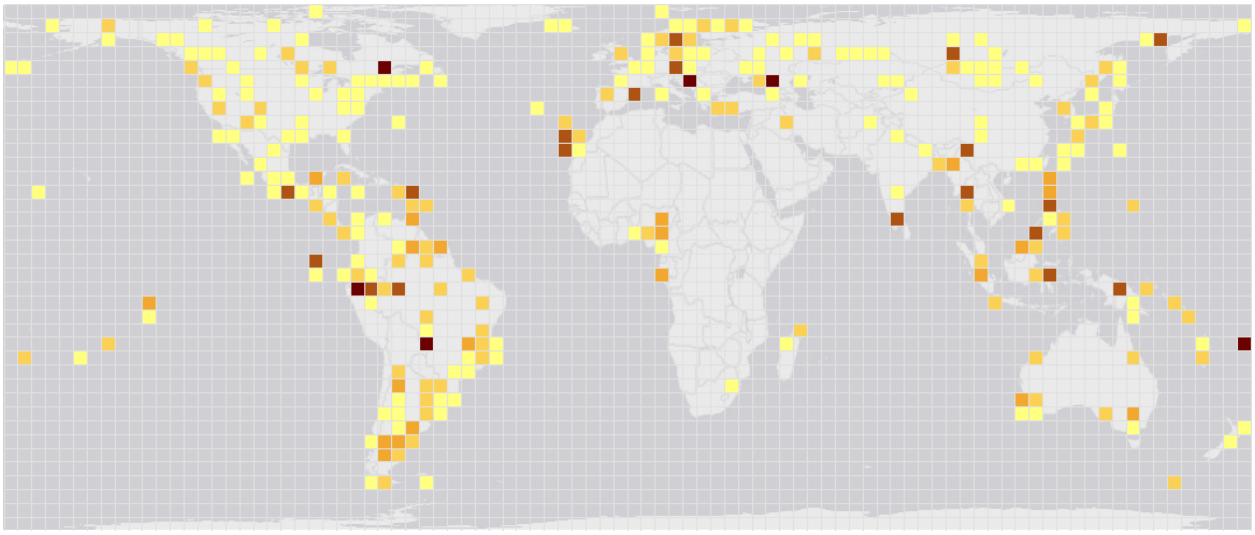

Figure 8: The figure shows the global distribution of genetic diversity using only the georeferenced sequences falling inside the breeding range. The map has a grid cell resolution of 4x4 and darker colours mean higher genetic diversity.

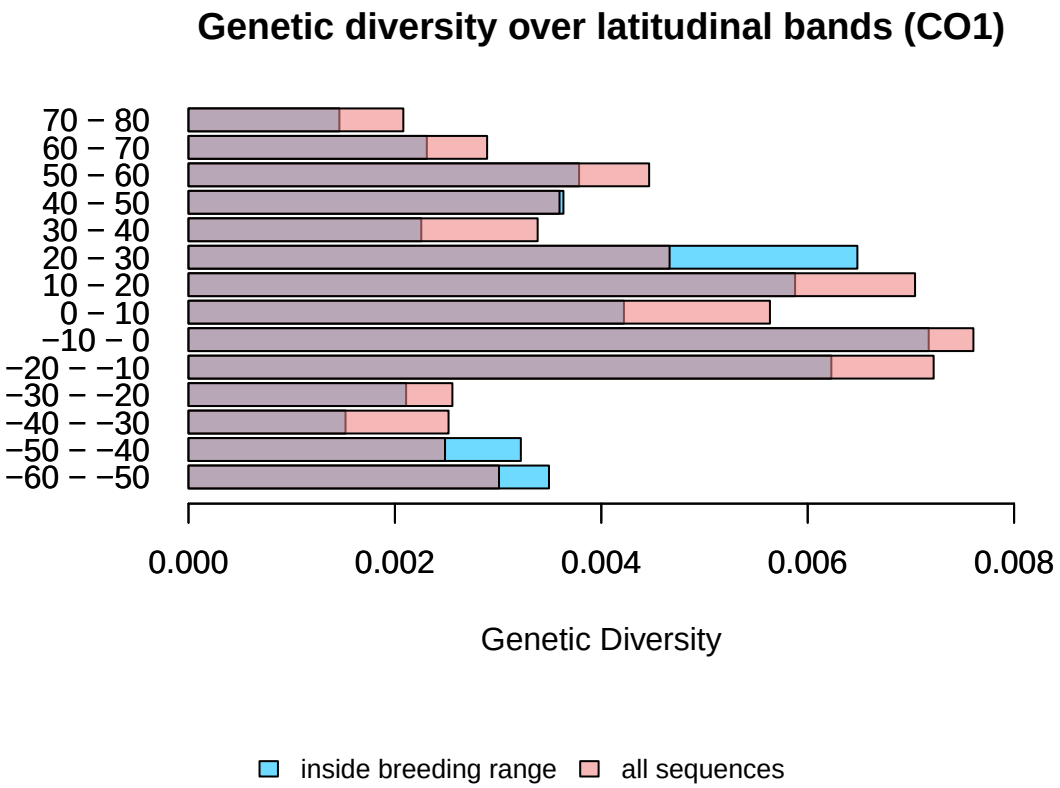

Figure 9: The figure shows the sovrapposition of results per latitudinal band obtained from the two different datasets.

CORRELATION ANALYSIS

In order to be observe wether using only the sequences inside the breeding range would affect the results, we also decided to run a correlation analysis that compares the two datasets for each loci.
For cytochrome-b, we obtained a correlation coefficient $R^2 = 0.85$; while for CO1, we obtained $R^2 = 0.80$. With these values we can say that the GD for both datasets varies in the same way; however, we are still uncertain about the "non-independence" of our GD values when doing the analysis, because one dataset is a subset of the other. This must be evaluated in order to assure the accuracy of the correlation.
Here are presented the two scatterplots, one for cytochrome-b and one for CO1, with the correlation resuls.

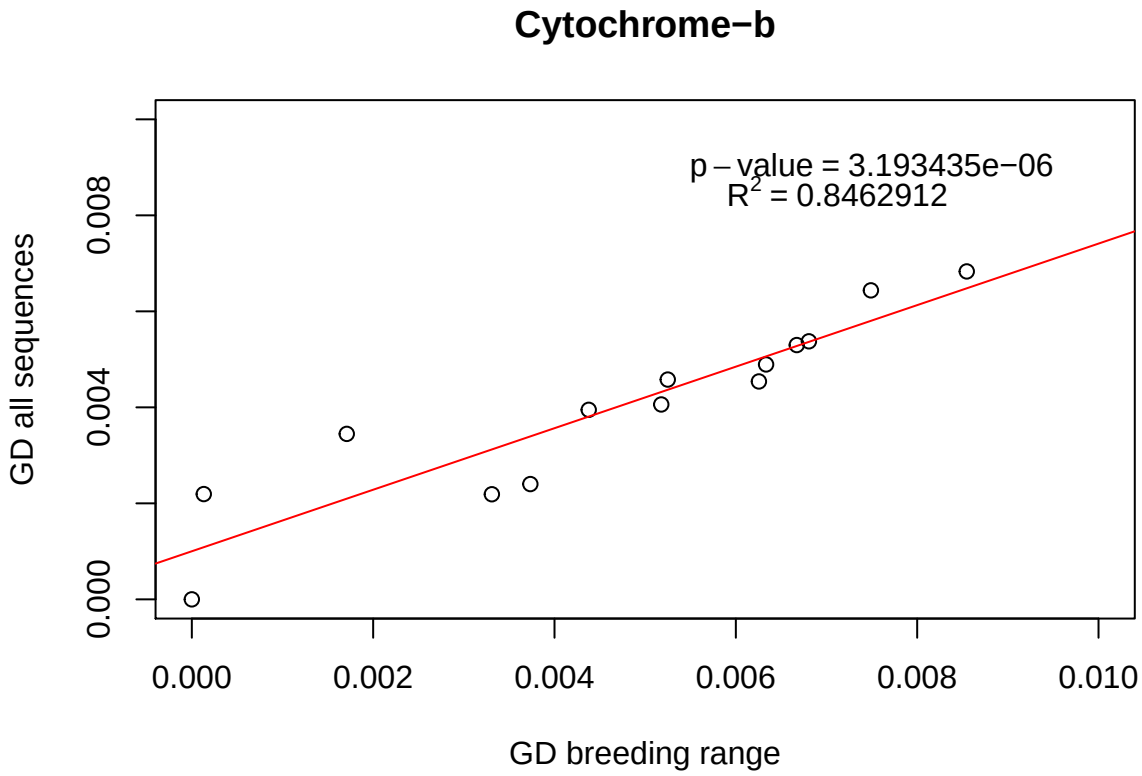

Figure 10: Regression analysis plot with the two cytochrome-b datasets on the two axes.

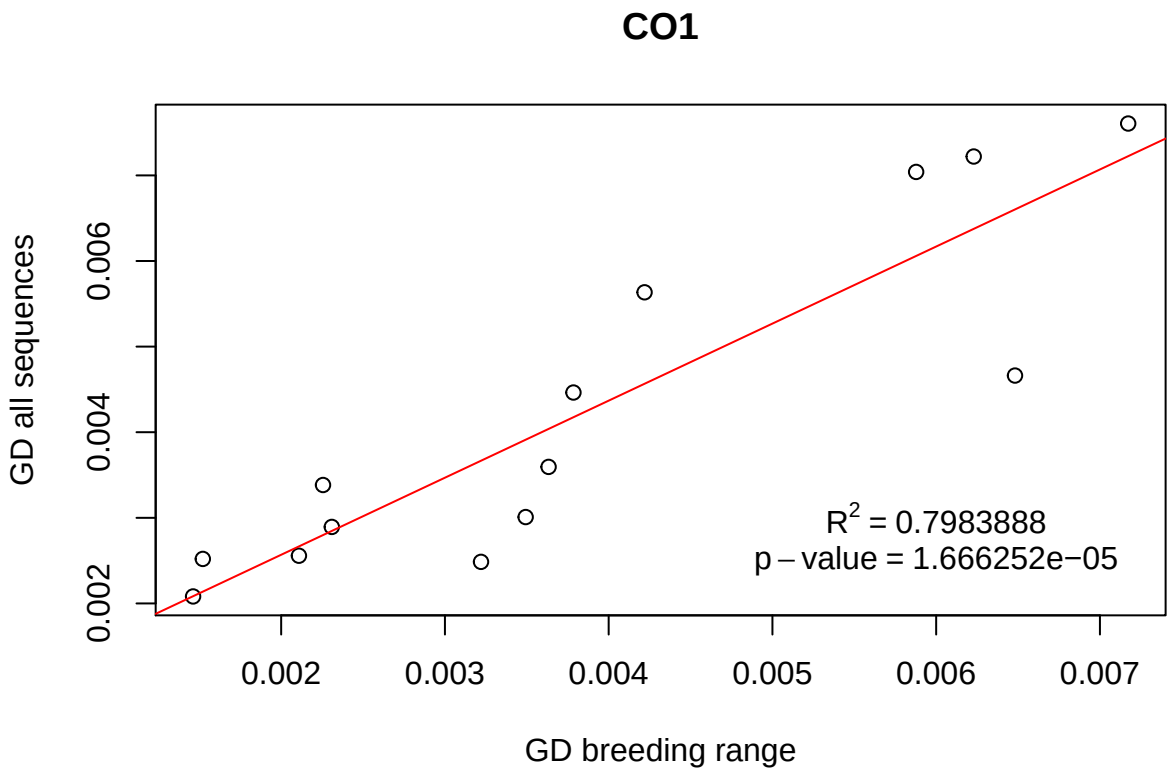

Figure 11: Regression analysis plot with the two CO1 datasets on the two axes.