

# Hawaiian Drosophila Investigation

---

SABRINA SAIPHOO



(Magnacca, 2010)

# Introduction

---

Hawaiian *Drosophila* have been widely investigated over the years due to their species diversity.

Due to the islands being isolated this poses a unique opportunity for genetic divergence; there are no predators and have a contained environment (Craddock, 1987).

This is also why these species strongly exhibit adaptive radiation, as there are a plethora of new species all descending from a common ancestor (Zimmerman, 1970).

This investigation will take a closer look at the species richness through analyzing unique BIN counts from the BOLD database.

# Investigation + Hypothesis

---

In this investigation, there will be a comparison between the distinct BIN counts for the Hawaii Islands and the mainland USA. For our purposes, mainland USA is considered to include everything except the Hawaiian Islands.

Due to the overwhelming research on adaptive radiation in Hawaii (Craddock, 1987), it is obviously predicted that the Hawaiian *Drosophila* will have a higher BIN diversity when compared to the mainland USA.

Next, we will investigate more specifically the distinct BINs found per island age, to determine whether older islands have greater BIN diversity than younger ones.

While the results could go either way, the initial hypothesis is that the older islands will have more BIN diversity, because there has been more time for colonization. Especially because the age of the islands range of 0.5 million to 5.1 million, with some differing by over 4.5 million years (Harrington, n.d) providing substantial time for evolutionary divergence.

There were around 36,000 records taken from BOLD

Cleaned data by removing missing values and duplicate entries for required columns.




Separated dataset into Hawai'i records and remaining U.S. records (Mainland USA).

Created a bar plot to compare distinct BIN counts for Hawaii, Mainland USA, and shared BINs.

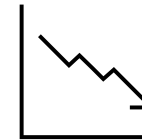


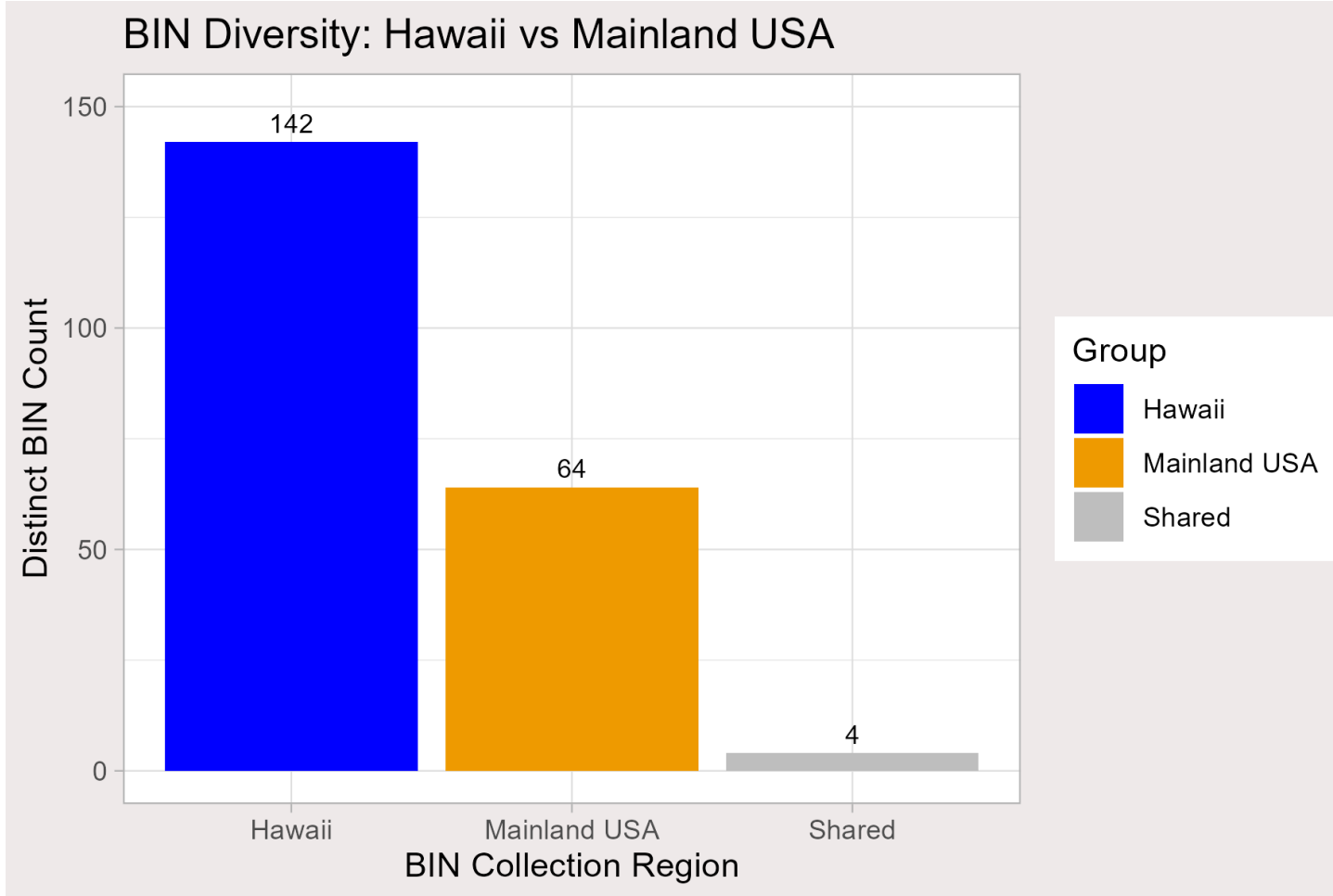
# Methods



Plotted accumulation curves to show if there are more BINs to be found in these regions.

Investigated island age and distinct BIN counts using a Negative Binomial regression





# Results 1

---

*The dataset contained 208 records from Hawaii and 1229 from the Mainland USA.*

# Hawaii exhibits significantly higher BIN diversity than the Mainland USA

---

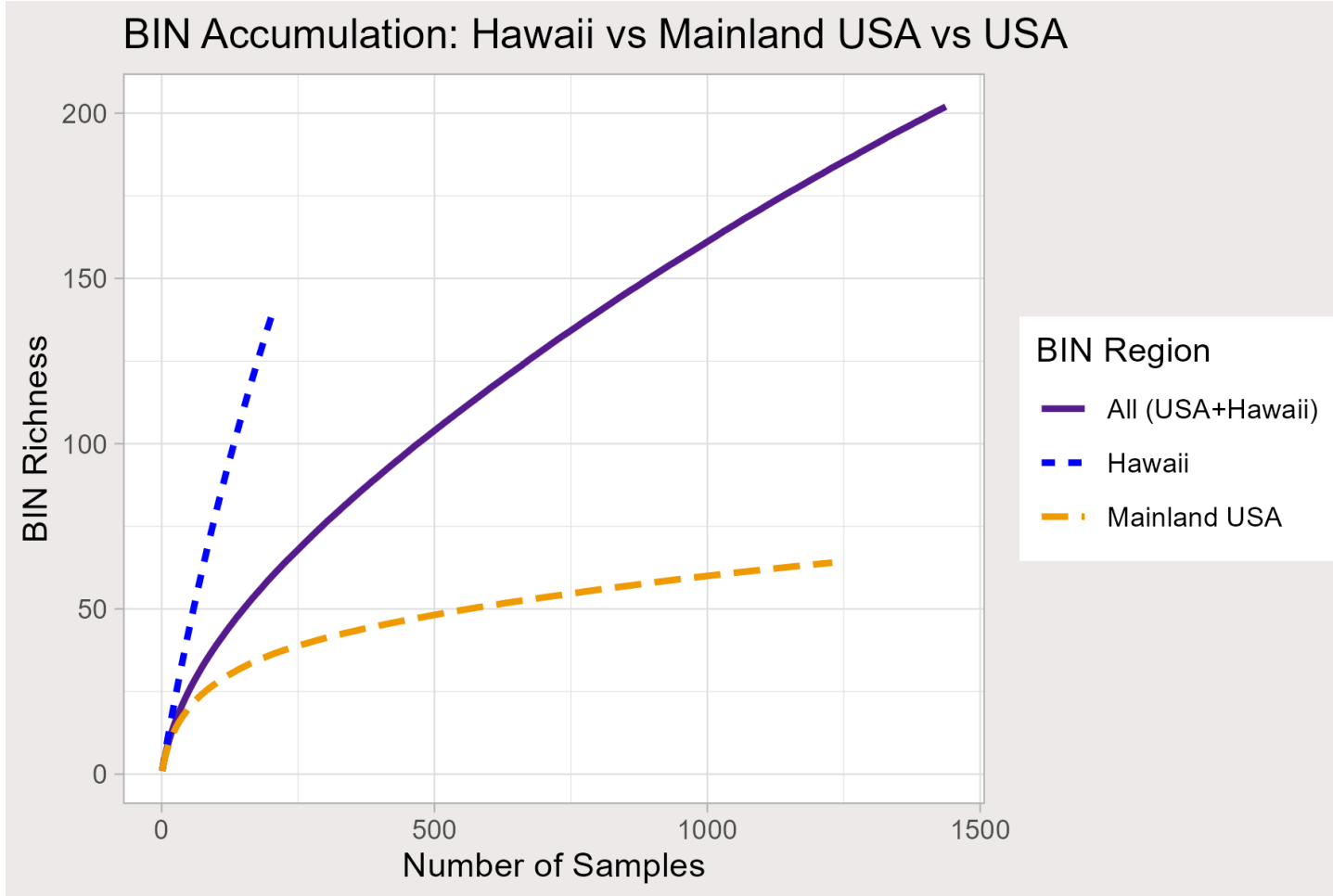
When evaluating the bar plot, it is evident that the distinct BIN count in Hawaii is more than twice that of the mainland USA. It can be inferred that Hawaii has a greater BIN diversity when compared to Mainland regions.

This is especially interesting given the difference in sampling effort, only 208 records from Hawai'i compared to 1,229 from the rest of the USA.

Out of these, 142 of 208 Hawaiian records were distinct BINs, versus 64 of 1,229 from the Mainland, with only four BINs shared between regions.

This pattern suggests that even with fewer total records, the rate of unique BIN discovery in Hawaii is disproportionately high, consistent with expectations for adaptive radiation in isolated environments (Craddock, 1987; Zimmerman, 1970).

While the pattern is evident, it is important to note that differences in sampling effort can also impact our results. Therefore, while the trend is strong, additional sampling is encouraged to confirm these findings.



# Results 2

---

# Species Accumulation in Hawaii remains increasing

---

## **Hawaii**

It is very evident from the species accumulation curve that the species in Hawaii are still constantly increasing. The increasing number of species relates to the concept of adaptive radiation, with many new species continuing to diversify in isolation (Craddock, 1987). It also demonstrates how geographic isolation can promote genetic divergence, unlike regions that are not geographically isolated (Craddock, 1987; Zimmerman, 1970).

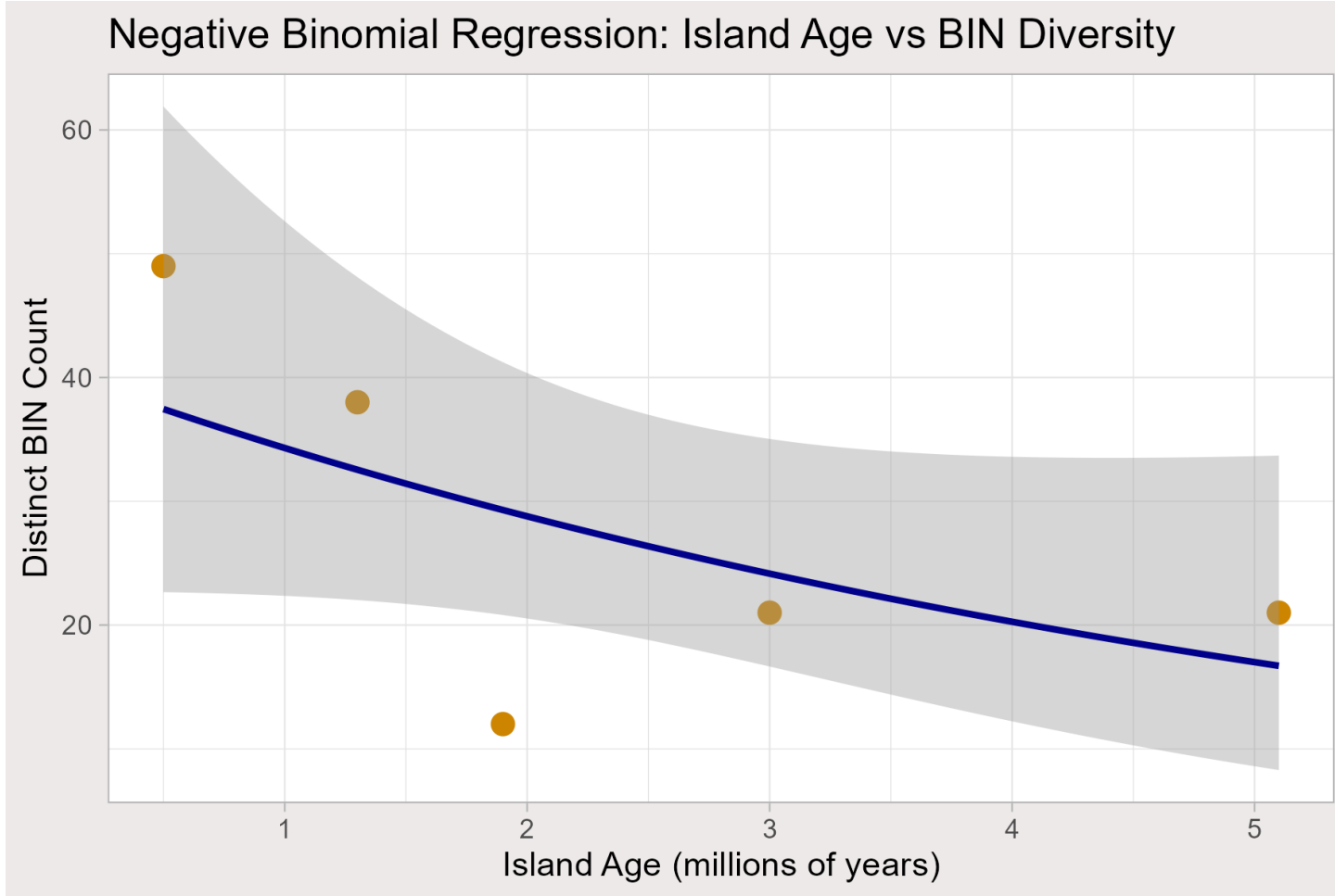
## **Mainland USA**

Since the Mainland USA includes samples from a wide range of environments and its curve is plateauing, we can conclude that it is closer to being fully sampled. If additional regions were included, we might observe an even stronger plateau. When comparing both regions, it becomes clear that the overall curve is driven largely by the Hawaiian data.

## **Validity**

Naturally, the difference in sample size must be considered (1,200 samples vs. 200 samples), as this may influence the trend. However, the steepness of the Hawaiian curve appears to reflect a biological signal rather than just under sampling, as it exceeds what would be expected by chance alone. The Hawaii region should continue to be sampled in order to fully confirm this trend.





# Results 3

# BIN richness decreases with Island Age

---

A **Negative Binomial regression** was chosen due to overdispersion in the data (Clapham, 2021), where the mean was much smaller than the variance. A slight negative trend was observed, showing that BIN richness decreases as island age increases. The coefficient ( $\beta = -0.175$ ) and p-value ( $p = 0.111$ ) indicate a negative relationship that isn't statistically significant but still noticeable.

As the Hawaiian Islands formed sequentially from Kauai to Hawaii, **Drosophila from older islands colonized the younger islands** (Hembry, 2021). This pattern aligns with the geological timeline, where older islands contain ancestral species to the species found on the newer islands (Hembry, 2021).

These findings reflect the ongoing process of **adaptive radiation**, where isolation and ecological opportunity support new species on younger islands, while older islands experience a gradual loss in diversity due to island erosion and species extinction (Craddock, 1987; Hembry, 2021; Zimmerman, 1970).

Although my hypothesis wasn't supported, the results suggest that **older islands act as source regions, providing the ancestral species that diversify on younger islands**. This could explain why younger islands show higher BIN counts. There's also a difference in sampling effort among islands, so additional data could help confirm this trend. When comparing BINs among islands, most were unique to their age range, showing little overlap overall.

# Conclusion

---

**BIN Diversity:** Younger Hawaiian Islands > Older Islands > Mainland USA

Hawaii is such an interesting location to do this kind of work because of its isolation aspect and continuous geological formation of new islands. Even within the scope of my small-scale analysis, the strength of adaptive radiation in *Drosophila* is clearly visible.

Although the dataset was limited, it was interesting to see consistent patterns emerge despite the small sample size. This highlights the importance of continued and standardized sampling efforts to further support these findings.

Each island hosts a unique array of *Drosophila* species, and these species often go on to colonize newly formed islands, continuing the cycle of diversification. This work reinforces how the Hawaiian Islands serve as an excellent model for studying evolution in real time.

# Acknowledgments

---

I had an initial conversation with Karl in developing my idea of investigating *Drosophila*. He posed the idea of comparing the diversity among island ages. Sharon Tsukernik also helped with the initial concept by giving me a better understanding of the different classes and helping me navigate the BOLD site.

Eva Innocente helped me understand different functions to add to the ggplots in order to format them the way I wanted. Farah Sadoon helped me realize that a linear regression wasn't appropriate for my count data, which led me to look into the negative binomial regression instead. Lastly, Liona ran my code and noticed I had a warning stemming from my species accumulation curve; this was a helpful note as I was able to modify the code to remove the warning.

# References

---

baptiste. (2016, March 16). *Manually colouring plots with scale\_fill\_manual in ggplot2 not working* [StackOverflow post]. StackOverflow. <https://stackoverflow.com/questions/36048033/manually-colouring-plots-with-scale-fill-manual-in-ggplot2-not-working>

Clapham, M. (2021, February 15). *Regression with Count Data: Poisson and Negative Binomial*. YouTube. <https://www.youtube.com/watch?v=uGKnoAw-PFQ>

Craddock, E. M. (1987). Degrees of reproductive isolation between closely related species of Hawaiian *Drosophila*. In M. J. D. White (Ed.), *Genetic mechanisms of speciation in insects*. Springer, Dordrecht. [https://doi.org/10.1007/978-94-010-2248-4\\_9](https://doi.org/10.1007/978-94-010-2248-4_9)

Frankham, R. (1997). Do island populations have less genetic variation than mainland populations? *Heredity*, 78(3), 311–327. <https://doi.org/10.1038/hdy.1997.46>

Harrington, D. (n.d.). *Hawaiian encyclopedia*. In *Hawaiian Encyclopedia: A Comprehensive Guide to the Hawaiian Islands—History, Culture, Native Species, Science*. Retrieved October 11, 2025, from <https://hawaiianencyclopedia.com/>

Hembry, D. H., Bennett, G., Bess, E., Cooper, I., Jordan, S., Liebherr, J., Magnacca, K. N., Percy, D. M., Polhemus, D. A., Rubinoff, D., Shaw, K. L., & O’Grady, P. M. (2021). Insect radiations on islands: Biogeographic pattern and evolutionary process in Hawaiian insects. *The Quarterly Review of Biology*, 96(4), 247–296. <https://doi.org/10.1086/717697>

Magnacca, K. (2010). *Drosophila silvestris* Kilohana. *Wikimedia Commons*. Retrieved from [https://commons.wikimedia.org/wiki/File:Drosophila\\_silvestris\\_Kilohana\\_5161a.jpg](https://commons.wikimedia.org/wiki/File:Drosophila_silvestris_Kilohana_5161a.jpg)

Ratnasingham, S., & Hebert, P. D. N. (2007). *Drosophila — Taxon page*. *Barcode of Life Data System (BOLD)*. Retrieved October 11, 2025, from [https://bench.boldsystems.org/index.php/Taxbrowser\\_Taxonpage?taxid=5757](https://bench.boldsystems.org/index.php/Taxbrowser_Taxonpage?taxid=5757)

The R Foundation / tidyverse. (n.d.). *FAQ: Customising in ggplot2: How can I change the background colour of plot?* In *ggplot2* (Version 3.x) [Online documentation]. Retrieved October 11, 2025, from <https://ggplot2.tidyverse.org/articles/faq-customising.html?q=background%20color#how-can-i-change-the-background-colour-of-plot>

Zimmerman, E. C. (1970). Adaptive radiation in Hawaii with special reference to insects. *Biotropica*, 2(1), 32–38. <https://doi.org/10.2307/2989786>