Global distribution of flower colors

Research proposal – The Center for AI and Data Science at Tel Aviv University (TAD)
Full grant

PIs: Yuval Sapir and Itay Mayrose
School of Plant Sciences and Food Security
Faculty of Life Sciences, Tel Aviv University

Abstract

Flower color is a critical trait influencing plant-pollinator interactions, stress tolerance, and horticulture value. Despite its importance, global flower-color distribution patterns, and the underlying drivers, remain poorly understood due to methodological constraints. The proposed research leverages recent technological advances and the increasing availability of citizen science data to encode the colors of all documented flowering plant species using automated image analysis, thereby presenting the first comprehensive global investigation of flower color distribution. Initially, we will validate our methods on a dataset of Israeli plant species with expert-validated color classifications. Subsequently, we will extend our analysis to images from the iNaturalist platform, optimizing our computational methods for variable image quality. Finally, we will process millions of geo-referenced images to analyze flower color distribution across all angiosperm species. Our data analysis and learning pipeline will accurately extract and classify flower colors, generating a global map of flower color distribution. We will employ multivariable regression models and structural equation modeling to decipher the ecological and climatic determinants of these patterns. This unprecedented global analysis will provide crucial insights into plant-pollinator relationships, predict responses to environmental change, and inform conservation strategies, addressing long-standing questions about flower color evolution at previously unattainable spatial and taxonomic scales.

<u>Research description</u>

Scientific background

Flower color is a central plant trait affecting plant-pollinator interactions, tolerance to abiotic stress, and of great commercial importance in the horticultural industry [1, 2]. Multiple factors have been hypothesized to affect floral color, spanning biotic interactions, environmental conditions, and evolutionary history. As ~87% of the flowering plant species are pollinated by animals [3], the prevailing hypothesis is that flower colors serve as adaptations to pollinators' visual perception, which, in turn, drives transitions in floral colors through pollinator-mediated selection [4-6]. Abiotic factors also impose selection on flower colors, with certain colors favored under certain environmental conditions and stress types [7-11]. Higher pigment concentration, hence darker flower colors, may provide better UV radiation protection in high-altitudes or latitude exposed to intense sunlight, while lighter flower colors (white and pale yellow) could reduce heat absorption in hot, arid environments [12-14]. Flower color may also be influenced by the evolutionary history of the floral population, as well as neutral processes [11, 15, 16].

The variability of floral coloration within an ecosystem is also affected by both biotic and abiotic factors. For example, flower colors in high-altitude regions in Taiwan show reduced diversity compared to low altitudes, likely due to intense selection from UV radiation and other environmental stressors [17]. In the Malesia archipelago, drought-prone seasonally dry regions exhibit increased frequencies of green and purple flowered species, suggesting adaptation to water stress [14]. Similarly, populations of *Anemone coronaria* in Israel maintain monomorphic red flowers in arid and desert regions, while in Mediterranean climates multiple flower colors are present (Fig. 1). These patterns suggest that environmental filtering plays a crucial role in determining regional flower color distributions.

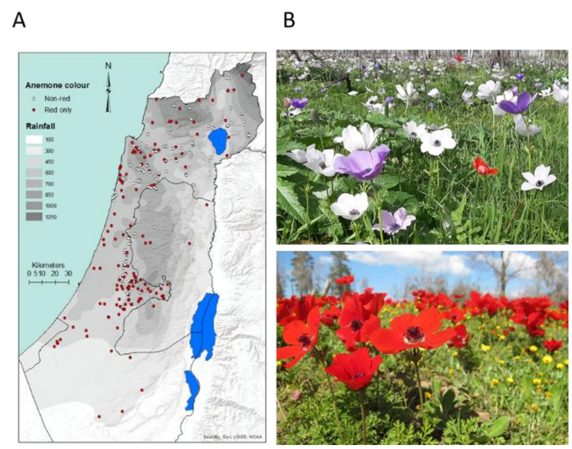

Figure 1 – A. Citizen science-based map (left) for flower colour of *A. coronaria* in Israel, showing that in the arid regions below 450 mm annual rain, red monomorphic populations (B, bottom) are exclusive, while in the north, Mediterranean climate, colour polymorphic populations are present (B, top). Y. Sapir, unpublished.

Previous studies regarding the distribution of flower color were limited in both taxonomic and geographic scope, although covering various regions and plant groups [11, 14, 17-23]; their restricted focus has, however, led to seemingly contradictory findings. Thus, for example, ecological hypotheses that predicted relationships between flower color and environmental factors in Eastern Australia were not supported in the regional scale [18], suggesting that broader geographic investigation is necessary. A major barrier to large-scale analyses has been the challenging process of color encoding, which traditionally relies on human visual inspection or manual reflectance measurements in the field, limiting our ability to generalize the effect of abiotic and biotic factors on flower color distribution.

Despite the fundamental importance of floral colors in shaping ecological and evolutionary processes, our understanding of its global distribution patterns and drivers remains limited by methodological constraints. Recent technological advances and the proliferation of citizen science platforms now provide an unprecedented opportunity to overcome these limitations. Leveraging millions of geo-referenced plant images available in digital repositories, combined with emerging learning approaches for automated image analysis, we propose to conduct the first comprehensive global investigation of flower color distribution. Our **aims** are: (1) To encode the colors of all documented flowering plant species (angiosperms) using automated image analysis; (2) To determine the distribution of flower colors across the globe; and (3) To decipher the ecological and climatic determinants underlying the global flower color distribution. This approach will allow us to address long-standing questions about flower color evolution at previously unattainable spatial and taxonomic scales.

Methodology

We propose to develop a general computational strategy to encode the flower color of all flowering plants with available image data, obtained through existing online repositories. Our research will proceed in three phases of increasing scope and complexity, each designed to validate and refine the computational predictions. The first phase will focus on a dataset of approximately 2,800 Israeli plant species, for which high-quality standardized images are available through the Flora of Israel website [24] and the Jerusalem Botanical Garden's photographic collection. For nearly all of these species, expert-validated color classification to discrete color bins is available and will served as ground-truth data for testing the automated classification procedure. In the second phase, we will extend our analysis to the same set of Israeli species, but using images from the iNaturalist citizen-science platform (www.inaturalist.org/). This phase will allow us to evaluate the developed procedures against the more variable image quality typical of citizen science data, where each species is represented by multiple images taken under different conditions. To ensure robust calibration, we will supplement our training data by manually encoding flower colors for 1,000 randomly selected iNaturalist images, helping us optimize our computational methods for the image quality typically found in citizen science repositories. Finally, we will expand our analysis to encompass all angiosperm species (>400,000; https://powo.science.kew.org/), processing millions of geo-referenced images from iNaturalist. This comprehensive approach will enable the first global-scale analysis of flower color distribution patterns.

We will employ a multi-stage deep learning pipeline to accurately extract and classify flower colors from digital images. The first stage involves flower detection and segmentation using Mask R-CNN architecture [25]. The model will be initialized with weights pre-trained on the COCO dataset [26] and fine-tuned on our manually annotated subset of flower images, ensuring precise isolation of flower regions from backgrounds, leaves, and other non-flower elements that could bias color analysis. Following segmentation, we will analyze the isolated flower regions using a custom convolutional neural network architecture designed to extract dominant color patterns while accounting for variations in lighting and exposure. This network will classify colors into standardized categories aligned with expert botanical classifications, handle multi-colored flowers and complex patterns such as spots, stripes, and gradients, and generate confidence scores for color classifications. The system will be trained progressively using our three-phase approach, with each phase incorporating additional validation data to improve accuracy and robustness. We will use classification confidence scores to flag potentially problematic classifications for manual review. In addition, the color identification of all images belonging to a single plant species will be consolidated to identify the consensus flower color of all plant species and to identify color polymorphic plant species for further analysis.

The classified flower colors for thousands of plant species, combined with millions of georeferenced occurrences, will enable us to create the first comprehensive global map of flower color distribution and analyze its drivers. Our analytical approach will follow the computational framework successfully implemented in our previous global-scale studies, where we revealed worldwide patterns of polyploid plant distribution [27] and plant life cycle evolution [28]. Briefly, we will obtain occurrence data for all angiosperm species from the Global Biodiversity Information Facility (GBIF; www.gbif.org), apply thorough taxonomic name resolution, and filter out unreliable records based on spatial precision and sampling methodology. The cleaned occurrence data will be mapped across the world's 814 ecoregions to calculate the proportional representation of different flower colors within each region. We will then analyze these distributions using multivariable regression models and structural equation modeling to test relationships with abiotic factors (including 19 BIOCLIM variables and solar radiation) and biotic factors (including species richness and global bee abundance; [29]). This unprecedented global analysis will reveal how evolutionary and ecological processes have shaped flower color diversity across the planet, providing crucial insights for understanding plant-pollinator relationships, predicting responses to environmental change, and informing conservation strategies.

Expected impact

Our proposed research is expected to have significant impact on both basic science and practical applications. By leveraging advanced image analysis techniques and extensive citizen science data, we will provide the first comprehensive global map of flower color distribution. This will deepen our understanding of how ecological and climatic factors shape floral traits, shedding new light on plant-pollinator interactions and evolutionary processes. The findings will be crucial for predicting plant responses to environmental changes, such as climate change and habitat loss, thereby informing conservation strategies and helping to preserve biodiversity. Beyond ecological insights, the development of automated image analysis methods for flower color classification will enhance botanical and agricultural research, improving the efficiency and accuracy of data collection and analysis. The integration of citizen science data will underscore the value of public participation in scientific research, fostering greater engagement in biodiversity monitoring efforts. Overall, this research will bridge critical knowledge gaps in plant ecology and evolution, providing a robust framework for future studies. It will also support conservation efforts by identifying regions and species most vulnerable to environmental changes, ultimately contributing to the preservation of global floral diversity.

References

1. van der Kooi, C.J., et al., *Evolution of insect color vision: from spectral sensitivity to visual ecology.* Annual review of entomology, 2021. **66**: p. 435-461.
2. Grünig, N., J.M. Horz, and B. Pucker, *Diversity and Ecological Functions of Anthocyanins.* Preprints, 2024.
3. Ollerton, J., R. Winfree, and S. Tarrant, *How many flowering plants are pollinated by animals?* Oikos, 2011. **120**(3): p. 321-326.
4. Rausher, M., *Evolutionary Transitions in Floral Color.* International Journal of Plant Sciences, 2008. **169**(1): p. 7-21.
5. van der Kooi, C.J., et al., *Competition for pollinators and intra-communal spectral dissimilarity of flowers.* Plant Biology, 2016. **18**(1): p. 56-62.
6. Bradshaw, H.D. and D.W. Schemske, *Allele substitution at a flower colour locus produces a pollinator shift in monkeyflowers.* Nature, 2003. **426**: p. 176 - 178.
7. Chalker-Scott, L., *Environmental Significance of Anthocyanins in Plant Stress Responses.* Photochemistry and Photobiology, 1999. **70**(1): p. 1-9.
8. Arista, M., et al., *Abiotic factors may explain the geographical distribution of flower colour morphs and the maintenance of colour polymorphism in the scarlet pimpernel.* Journal of Ecology, 2013. **101**(6): p. 1613-1622.
9. Vaidya, P., et al., *Ecological causes and consequences of flower color polymorphism in a self-pollinating plant (Boechera stricta).* New Phytologist, 2018: p. doi: 10.1111/nph.14998.
10. Dick, C.A., et al., *Arctic Mustard Flower Color Polymorphism Controlled by Petal-Specific Downregulation at the Threshold of the Anthocyanin Biosynthetic Pathway.* PLoS ONE, 2011. **6**(4): p. e18230.
11. Wang, Q., et al., *Evolutionary history and climate conditions constrain the flower colours of woody plants in China.* Journal of Plant Ecology, 2021. **15**(1): p. 196-207.
12. Tripp, E.A., et al., *Evolutionary and Ecological Drivers of Plant Flavonoids Across A Large Latitudinal Gradient.* Molecular Phylogenetics and Evolution, 2018.

13. Berardi, A.E., et al., *Elevational divergence and clinal variation in floral color and leaf chemistry in Silene vulgaris.* American Journal of Botany, 2016. **103**(8): p. 1508-1523.
14. Trethowan, L.A., et al., *Island area, isolation and climate effects upon flower traits in a megadiverse archipelago.* Journal of Biogeography, 2024. **51**(9): p. 1659-1669.
15. Roguz, K., et al., *All the Colors of the Rainbow: Diversification of Flower Color and Intraspecific Color Variation in the Genus Iris.* Frontiers in Plant Science, 2020. **11**(1519): p. 569811.
16. Wang, H., et al., *Neutral processes contribute to patterns of spatial variation for flower colour in the Mediterranean Iris lutescens (Iridaceae).* Annals of Botany, 2016. **117**(6): p. 995-1007.
17. Tai, K.-C., et al., *Floral Color Diversity: How Are Signals Shaped by Elevational Gradient on the Tropical–Subtropical Mountainous Island of Taiwan?* Frontiers in Plant Science, 2020. **11**.
18. Dalrymple, R.L., et al., *Macroecological patterns in flower colour are shaped by both biotic and abiotic factors.* New Phytologist, 2020. **228**(6): p. 1972-1985.
19. Shrestha, M., et al., *Flower colour and phylogeny along an altitudinal gradient in the Himalayas of Nepal.* Journal of Ecology, 2014. **102**(1): p. 126-135.
20. de Camargo, M.G.G., et al., *How flower colour signals allure bees and hummingbirds: a community-level test of the bee avoidance hypothesis.* New Phytologist, 2019. **222**(2): p. 1112-1122.
21. Kantsa, A., et al., *Community-wide integration of floral colour and scent in a Mediterranean scrubland.* Nature Ecology & Evolution, 2017. **1**(10): p. 1502-1510.
22. Warren, J. and S. Mackenzie, *Why are all colour combinations not equally represented as flower-colour polymorphisms?* New Phytologist, 2001. **151**(1): p. 237-241.
23. León-Osper, M. and E. Narbona, *Unravelling the mystery of red flowers in the Mediterranean Basin: How to be conspicuous in a place dominated by hymenopteran pollinators.* Functional Ecology, 2022. **36**(11): p. 2774-2790.
24. Danin, A. and O. Fragman-Sapir. *Flora of Israel online, http://flora.org.il/plants/*. 2016+ 22 January, 2025].
25. He, K., et al., *Mask R-CNN.* Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2017: p. 2980-2988.
26. Lin, T.-Y., et al. *Microsoft coco: Common objects in context*. in *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13*. 2014. Springer.
27. Rice, A., et al., *The global biogeography of polyploid plants.* Nature Ecology & Evolution, 2019. **3**(2): p. 265-273.
28. Poppenwimer, T., I. Mayrose, and N. DeMalach, *Revising the global biogeography of annual and perennial plants.* Nature, 2023. **624**(7990): p. 109-114.
29. Orr, M.C., et al., *Global Patterns and Drivers of Bee Distribution.* Current Biology, 2021. **31**(3): p. 451-458.e4.

List of participants

CURRICULUM VITAE: YUVAL SAPIR

Yehuda Naftali Botanic Garden, School of Plant Sciences and Food Security, George S. Wise Faculty of Life Sciences, Tel Aviv University, Tel Aviv 69978

Phone (972) 54-7203140; e-mail: sapiry@tauex.tau.ac.il

Education

1999-2004	Ph.D., Evolution, Systematics and Ecology, The Hebrew University of Jerusalem
1997-1999	M.Sc., Evolution, Systematics and Ecology, The Hebrew University of Jerusalem
1994-1997	B.Sc., Life Sciences, The Hebrew University of Jerusalem

Employment

2012-present	Senior Lecturer, School of Plant Sciences and Food Security, Tel Aviv University
2009-present	Director, Tel Aviv University Botanic Garden.
2004-2009	Post-doctoral fellow, Indiana University, USA; University of Georgia, USA; Ben-Gurion University

External Grants

American Iris Society Foundation (2002, 2006, 2017), Ministry of Agriculture (2014), ISF (2016, 2022), ISF-FIRST (2016, 2018), DFG (2021), Google (2023).

Editorial Boards

Journal of Evolutionary Biology, American Journal of Botany, Journal of Pollination Ecology.

Supervision of graduate students and postdoctoral fellows

2009–Present: 6 post-docs, 6 Ph.D., and 16 M.Sc. students.

Selected relevant publications (Google Scholar H index 24. 1,768 citations)

Lebel M, Oboloski U, Hadany L, Sapir Y. 2018. Pollinator-mediated selection on floral size and tube color in *Linum pubescens*: Can differential behavior and preference in different times of the day maintain dimorphism? *Ecology and Evolution*; 8(2):1096-1106.

Veits M, Khait I, Obolski U, Zinger E, Boonman A, Goldshtein A, Saban K, Seltzer R, Ben-Dor U, Estlein P, Kabat A, Peretz D, Ratzersdorfer I, Krylov S, Chamovitz D, Sapir Y, Yovel Y, Hadany L. 2019. Flowers respond to pollinator sound within minutes by increasing nectar sugar concentration. *Ecology Letters* 22(9):1483-1492.

Roguz K, Gallagher MK, Senden E, Bar-Lev Y, Lebel M, Heliczer R, Sapir Y. 2020. All the Colors of the Rainbow: Diversification of Flower Color and Intraspecific Color Variation in the Genus *Iris*. *Frontiers in Plant Science*; 11:569811.

Sapir Y, Gallagher MK, Senden E. 2021. What maintains flower colour variation within populations? *Trends in Ecology and Evolution*; 36(6):507-519.

CURRICULUM VITAE: ITAY MAYROSE

School of Plant Sciences and Food Security, George S. Wise Faculty of Life Sciences,
The Edmond J. Safra Center for Bioinformatics
Tel Aviv University, Tel Aviv 69978

Phone (972) 3-6407212; FAX (972) 3-6409380; e-mail: itaymay@tauex.tau.ac.il

Education

2008-2011 Ph.D. (with distinction), Cell Research and Immunology, Tel-Aviv University

2002-2004 M.Sc., Zoology, Tel-Aviv University

1997-2001 B.Sc., Multidisciplinary program in Computer Science and Biology, Tel Aviv University

Employment

2021-to date Professor, School of Plant Sciences and Food Security, Tel Aviv University

2016-2021 Associate Professor, School of Plant Sciences and Food Security, Tel Aviv University

2011-2016 Senior lecturer, Department of Molecular Biology & Ecology of Plants, Tel Aviv University

2008-2011 Post-doctoral fellow, Biodiversity Research Institute, University of British Columbia

External Grants

ISF (2012-2016, 2017-2021, 2021-2026), BSF (2014-2016), BARD (2018-2021), Marie Curie (2011-2015), Ministry of Agriculture (2016-2019, 2020-2023), NSF-BSF (2017-2020, 2020-2024), Chief Scientist of Israel (2020-2023, 2023-2025).

Major positions at Tel Aviv University

2012–2023 Head of the M.Sc. track in Theoretical and Mathematical Biology

2021–Present Head of the B.Sc. track for excellent undergraduate students in the Life Sciences Faculty

Editorial Boards

New Phytologist, Methods in Ecology and Evolution, Genome Biology and Evolution

Supervision of graduate students and postdoctoral fellows

2011–Present 5 post-docs, 9 Ph.D., and 15 M.Sc. studetns all at Tel-Aviv University

Patents

2022 U.S. Provisional Patent Application No. 63/329,506. Systems and methods for genome-scale targeting of functional redundancy in plant.

Selected relevant publications (Google Scholar H index 40. 15,252 citations)

Azouri D, Granit O, Alburquerque M, Mansour Y, Pupko T, Mayrose I. The tree reconstruction game: phylogenetic reconstruction using reinforcement learning. *Molecular Biology and Evolution*. 41(6):msae105. 2024

Ecker N, Huchon D, Mansour Y, Mayrose I, and Pupko T. 2024. A machine-learning based alternative to phylogenetic bootstrap. Bioinforamtics. 40(Supplement_1):i208-i217

Poppenwimer T, Mayrose I, De-Malach N. Revising the global biogeography of annual and perennial plants. Nature. 624:109–114. 2023.

Azouri D, Abadi S, Mansour Y, Mayrose I, Pupko T. 2021. Harnessing machine learning to guide phylogenetic-tree search algorithms. *Nature communication*. 12:1983.

Abadi S, Avram O, Rosset S, Pupko T, Mayrose I. 2020. ModelTeller: model selection for optimal phylogenetic reconstruction using machine learning. *Molecular Biology and Evolution*. 37(11):3338–3352.

Abadi S, Azouri D, Pupko T, Mayrose I. 2019. Model selection may not be a mandatory step for phylogeny reconstruction. *Nature Communication*. 10:934.

Rice A, Smarda P, Novosolov M, Drori M, Glick L, Sabath N, Meiri S, Belmaker J, Mayrose I. 2019. The Global Biogeography of Polyploid Plants. *Nature Ecology and Evolution*. 3:265-273.

<u>A statement of ethical compliance</u>
Not relevant

Required budget

Item	Details	Year 1	Year 2
Personnel	MSc student (150% scholarship)	₪72,000	₪92,000
Computers	Laptop	₪12,000	
Cloud services		₪30,000	₪25,000
Total		₪114,000	₪117,000