

Supporting Information

Joppa et al. 10.1073/pnas.1109389108



Fig. S1. Original TDWG regions.



Fig. S2. Regions used to analyze numbers of missing species around the world. Original TDWG regions are delineated with gray outlines, and the regions actually analyzed are uniquely colored. Only TDWG regions that shared a border were clustered together for our analysis.

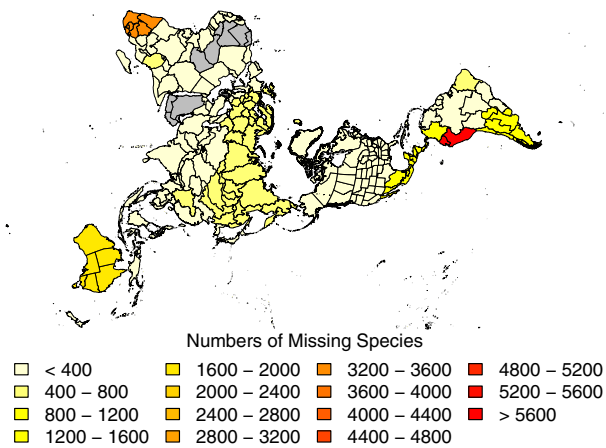


Fig. S3. The estimated raw numbers of species remaining to be described in each of the regions depicting in Fig. S2. Actual values can be found in Dataset S1, column 16.

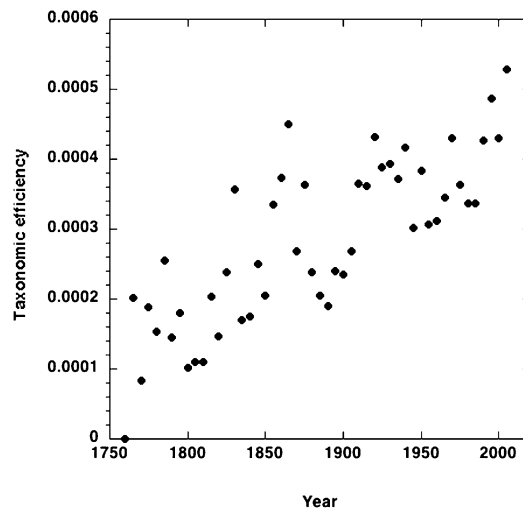


Fig. S4. Taxonomic efficiency, or the number of species described per taxonomist, corrected for the remaining number of missing species, $S_i / (T_i) * (S_T - \Sigma S_i)$, increases linearly with time in this example of approximately 50,000 flowering plants in families other than monocots. This example is typical of most comparable relationships for subsets of the flowering plants, although, for some, there is not such an obvious increase over time.

Dataset S1. Detailed summary statistics on model fit for all regions included in the analysis

[Dataset S1](#)

Column labeled Region Code Name provides a link between this dataset and [Dataset S2](#), which lists the component TDWG regions that comprise each clustered region. Column labeled Current Rank lists the rank from 1 to 50 of each region based on the current total number of endemic species, and Future Rank does the same based on MLE of the S_T parameter. Rank Change calculates change in position between current and future. We present number of species in each region in two different ways, as raw counts ("raw") and as a percent of the total ("percent"). The ratio of increase is calculated as the current total number of species in a region divided by the MLE-estimated total.

Dataset S2. Original TDWG information on each region, with a linking column to [Dataset S1](#) and the regions that were eventually clustered together for the final analysis

[Dataset S2](#)