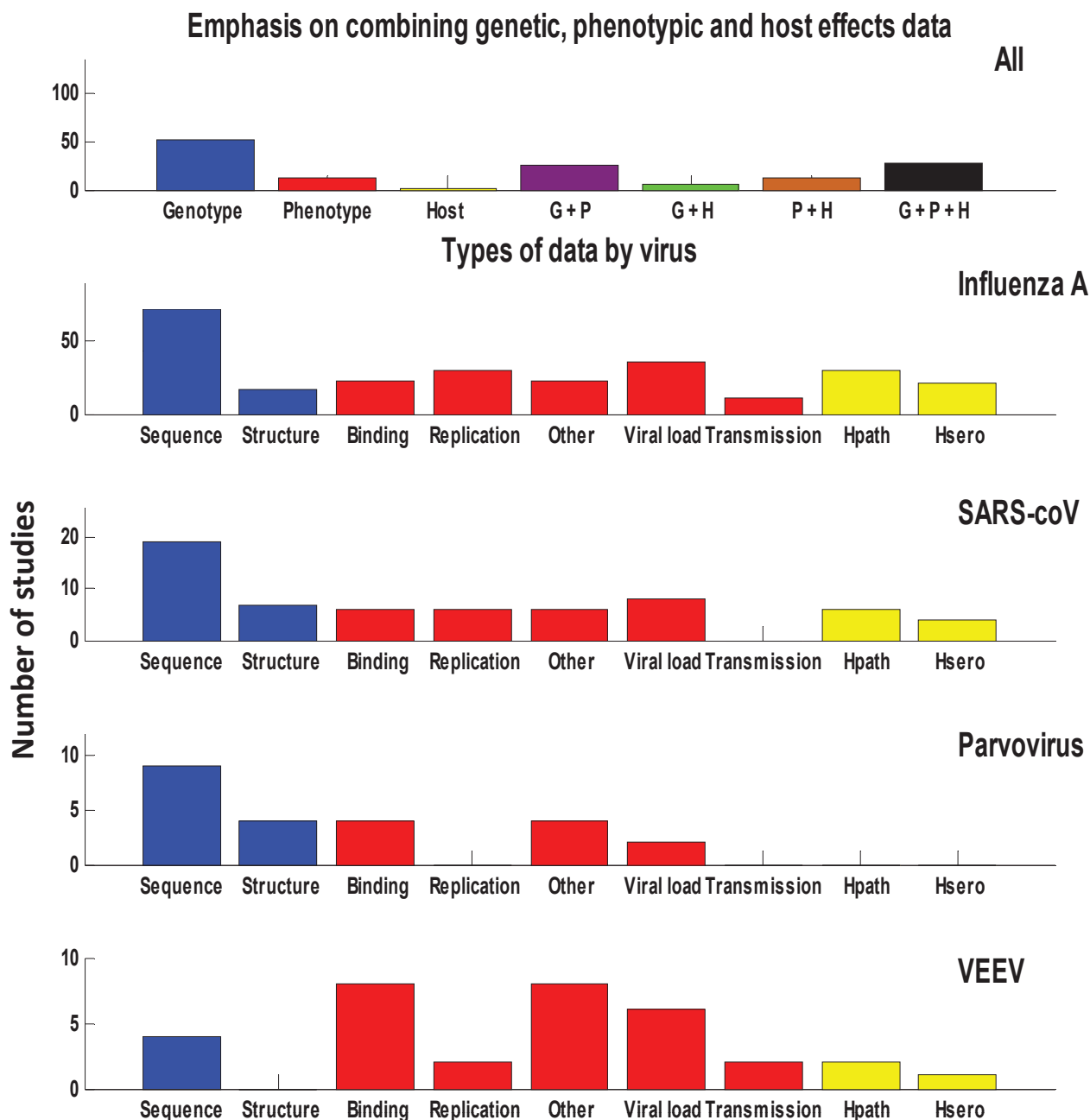


S5 | Types of data used for analyses of evolutionary processes.



Studies are grouped by whether they included genetic, phenotypic, host effects data or combinations of these categories (e.g. G+P means that the study includes genotypic and phenotypic data; top plot). Only studies that tested evolutionary hypotheses (black Genetic data (Genotype, blue) includes genetic sequence or protein structure. Phenotype (red) involves measurements from Fig. 1) are included. Measurement of a viral trait such as receptor binding, polymerase replication efficiency, withinhost fitness (time course of infectious particles), transmission rate or other. Host (yellow) involves measurement of a host pathology (Hpath) or immune response (Hsero). Categories in top plot are mutually exclusive while categories for each separate virus (second plot to last plot) are not. Genotype (G), Phenotype (P), Host (H). The maximum value for each of the Y-axes is the total number of papers surveyed (134 for All, top).