

Fig. 5 Capsid volume usage is often low and varies significantly among viral families. **a** A scatter plot demonstrating the volume usage (in %) with respect to genome lengths. Only the X-axis is in log scale. The ellipses were created by first calculating the volume usage percentage for each genus separately, and then drawing the families by the distributions of these values. The analysis covers all icosahedral viruses that are associated with detailed 3D information. There are 24 such icosahedral families: 1 – Partitiviridae, 2 – Tymoviridae, 3 – Dicistroviridae, 4 – Rudiviridae, 5 – Bromoviridae, 6 – Togaviridae, 7 – Tectiviridae, 8 – Reoviridae, 9 – Papillomavirida, 10 – Chrysoviridae, 11 – Circoviridae, 12 – Phycodnavirida, 13 – Tombusviridae, 14 – Bimaviridae, 15 – Cystoviridae, 16 – Caliciviridae, 17 – Hepadnaviridae, 18 – Totiviridae, 19 – Leviviridae, 20 – Nodaviridae, 21 – Adenoviridae, 22 – Flaviviridae, 23 – Polyomaviridae, 24 – Picornaviridae. Spearman's rank correlation is not significant: p = -0.17, p-value = 0.42. **b** An arbitrary sample of 10 families presented in (a), demonstrating the proportions of their capsid and genome sizes, from which the volume usage is derived. A single genus was chosen to represent each family, illustrating its capsid (with surface images from VIPERdb) and genome size (showing a bar proportional to its length that also displays the number of strands, and using the color of the relevant viral group). The radii of the capsid images are proportional to their outer radius (although it's the inner radius that determines the volume usage; both are written). Additional structural details (number of capsid subunits and T number) are also shown. The representative genus of each family was chosen by uniform rule - the one with the largest inner radius. This rule also applied for the displayed VIPERdb record