

Scatter plot showing the relationship between GC content (x-axis, 0 to 100%) and Tm (°C) (y-axis, 2400 to 2475) for 1000 random sequences. The plot includes individual data points (black dots) and a green line representing the mean Tm. The Tm values generally decrease as GC content increases, with a slight plateau around 50% GC.

A scatter plot showing the relationship between $\log_{10}(1 + 1000 \cdot (1 - R^2))$ on the x-axis and $\log_{10}(1 + 1000 \cdot (1 - R^2))$ on the y-axis. The data points are black dots, and a solid green line represents the identity line ($y = x$). The plot shows a strong positive correlation, with most points falling above the identity line, indicating that the y-axis values are generally higher than the x-axis values.

Scatter plot showing the relationship between the number of genomes in the pangenome (x-axis) and the number of genes in the core genome (y-axis). The x-axis ranges from 1 to 6, and the y-axis ranges from 2480 to 2560. A green line represents a linear regression fit, showing a positive correlation. Data points are black dots.

Number of genome(s) in the pangenome	Number of genes in the core genome
1	2470
1	2484
1	2485
1	2486
1	2490
2	2506
2	2508
2	2527
2	2527
2	2527
3	2516
3	2517
3	2539
3	2539
3	2544
4	2521
4	2527
4	2527
4	2543
4	2554
5	2551
5	2551
5	2560
5	2560
6	2562
6	2562
6	2562