

Supplemental Information

Table S1. Summary table showing the number of MPD datasets that measure responses using the indicated number of inbred strains. For our analyses, we selected 2435 MPD datasets that measured a response in 10 or more inbred strains. Of note, 43% of these datasets analyzed 27, 24, 23, or 20 inbred strains.

# of Strains	Number of Datasets
10	9
12	68
13	48
14	33
15	82
16	136
18	103
19	47
20	159
21	41
22	53
23	162
24	265
25	132
26	88
27	416
28	147
29	138
30	73
31	94
32	80
33	61

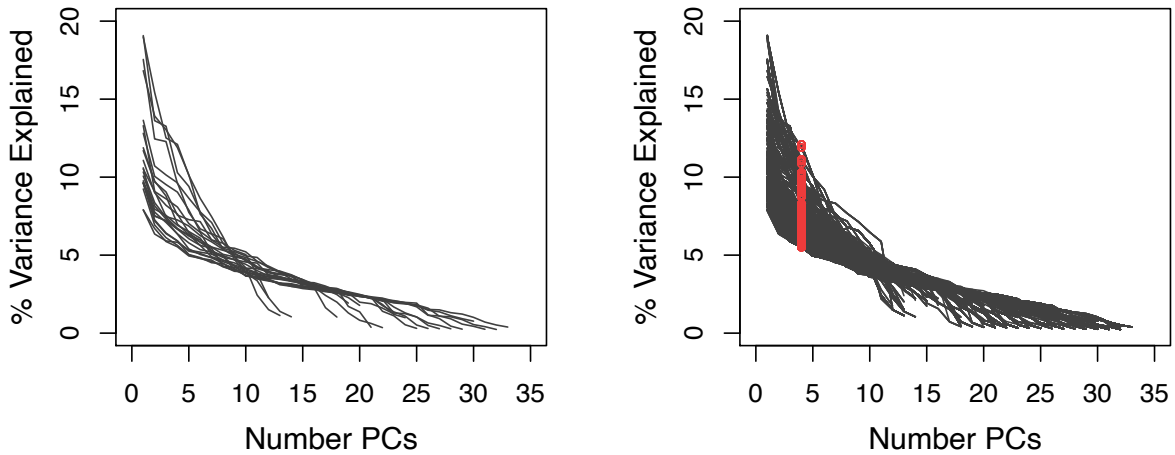


Figure S1. The percentage of variance that is explained when the indicated number of principal components (PC) are used to analyze 22 selected MPD datasets (*Left*) or all 2435 MPD datasets (with ≥ 10 inbred strains) (*Right*). Every MPD dataset analyzed a response in a panel of inbred strains (range 10 to 33 strains). Each individual curve represents a PCA for one mouse population for a dataset. The red circles show the percentage of the variance that is explained by the 4th PC, which ranges from 5.5%-12.1%. The total variance explained ranges between 26% and 59% when the first four PCs are used.

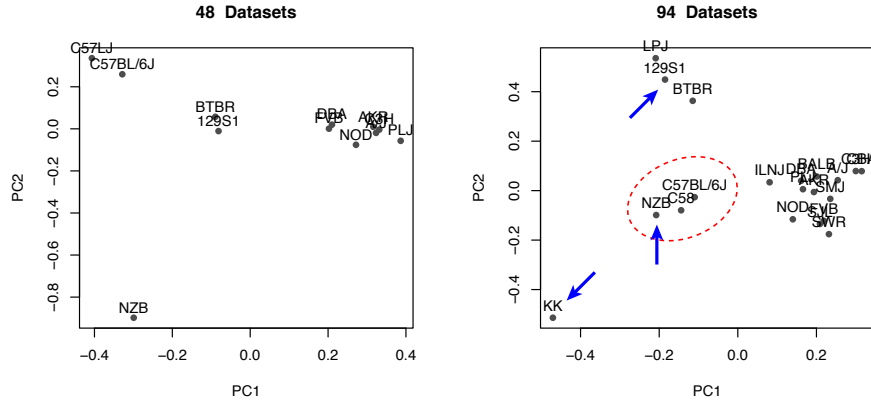


Figure S2. PCA plots (using the first two PCs) examine the genetic relationships of the 12 strains that are evaluated in 48 MPD datasets (*Left*) or the 20 inbred strains evaluated in 94 MPD datasets (*Right*). **Left:** In this graph, although NZB is separated from the other strains; a single strain cannot by itself be classified as a sub-population. Also, while there appear to be 4 sub-populations in this plot, three of the groups contain only 1 or two strains. **Right:** This graph appears to have 4 sub-groups of strains. However, KK (a global group 2 strain, blue arrow at bottom) forms a strain group that is separated from other group 2 strains (NZB, BTBR, 129S1). Moreover, global group 2 (NZB) and group 1 (C57BL/6J, C58) strains form one group in this graph (red circle). Overall, when a small number of strains are evaluated, the sub-population structure is highly variable, and it can be significantly altered by adding or deleting a single strain.

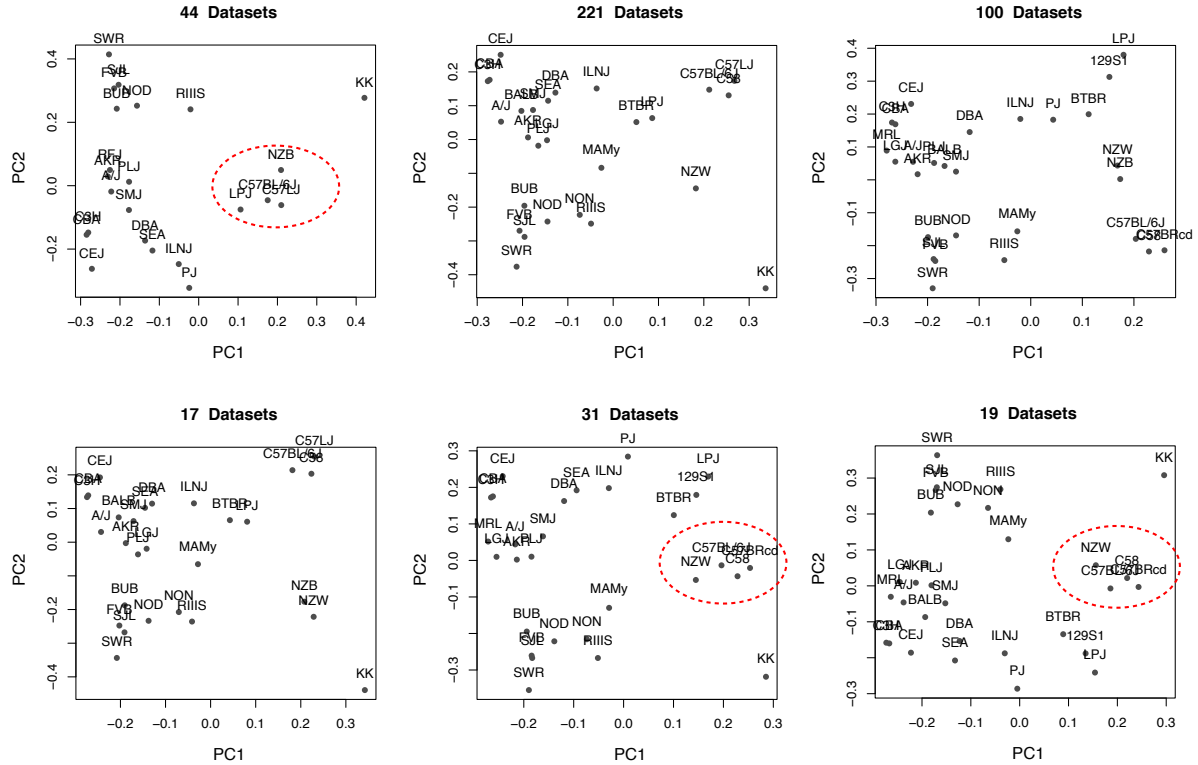


Figure S3. Six commonly used panels of inbred strains, whose responses were measured in 432 MPD datasets, lack population sub-structure. Scatter plots, which show the first two PCs, were generated to examine population structure among these commonly used strains. These MPD datasets measure responses (from top left to bottom right) in the same set of 23 strains (44 datasets), 27 strains (221 datasets); 28 strains (100 datasets) and a second set of 17 datasets used a different 28 strain panel; 29 strains (31 datasets); and 30 strains (19 datasets). The red circles indicate groupings where the local population sub-structure violates the global PS : group 1 (C57BL related) strains are intermixed group 2 (NZB, NZW, LPJ) strains.

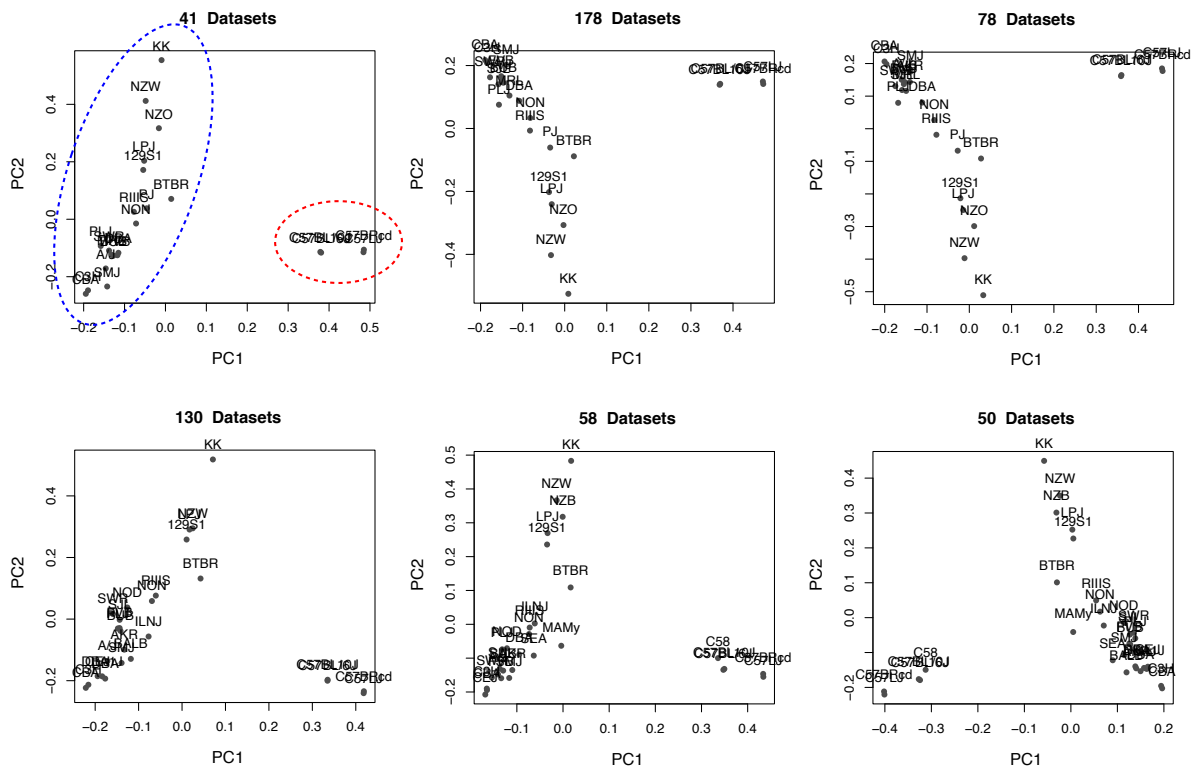


Figure S4. Six commonly used panels of inbred strains, which were used to generate 535 MPD datasets, have population sub-structure. Scatter plots showing the first two PCs were generated to examine population structure among the 23 to 32 inbred strains evaluated in these 535 MPD datasets. Responses in these datasets were measured (from top left to bottom right) using the same set of 23 strains (41 datasets), 24 strains (178 datasets); 25 strains (78 datasets); 27 strains (130 datasets), 29 strains (58 datasets); and 32 strains (50 datasets). The group 1 strains (C57BL related) are clearly separated from other strains in these graphs (red circle, top left panel). However, global group 2 and group 3 strains are broadly distributed in these graphs, and they cannot be separated into distinct sub-groups (see blue circle).