

| Table 1. Pseudo r2 values from multinomial regression of 5 first PC components of genomic PCA on various metadata features (features described in model column). ML r2 value reported in main text to test for significance of association between metadata features and genomic PCA dimensions. Last model (Isolation_country_conditioning_MLST) also includes MLST label as a predictor for Isolation Country metadata feature. | | | | | | | | |
|--|----------------|-----------|-----------------|-------------|-------------|-------------------------------------|--|--|
| Ilh | IlhNull | G2 | McFadden | r2ML | r2CU | Model | | |
| -8769.881 | -9749.848 | 1959.935 | 0.101 | 0.245 | 0.261 | Isolation_country | | |
| -13697.680 | -13822.948 | 250.536 | 0.009 | 0.049 | 0.049 | Collection_year | | |
| -3255.736 | -9301.580 | 12091.690 | 0.650 | 0.959 | 0.966 | MLST_type | | |
| -835.584 | -3723.990 | 5776.813 | 0.776 | 0.784 | 0.910 | Isolation_country_conditioning_MLST | | |