|  |  |  |
| --- | --- | --- |
| Author | Paper Title | Information Issues |
| Payen 2016 | High-Throughput Identification of Adaptive Mutations in Experimentally Evolved Yeast Populations | * Number of populations and clones listed in the paper no not match the number of populations and clones in the SI. * Populations and clones were not consistent, i.e a clone would be labeled as a population and given a frequency. * The references to different papers do not match the information of that paper, i.e when doing the c-hyper calculations, I do not get the same value for the “Lang” reference and the actual Lang paper. * Doesn’t explain which are populations and which are clones. |
| Kacar 2017 | Experimental Evolution of Escherichia coli Harboring an Ancient Translation Protein | * The number of mutations indicated in the paper does not match the number of mutations in the SI. * Refers to a Non-Hybrid lineage but has no information about it in the paper or SI. |
| McCloskey 2018 | Evolution of gene knockout strains of *E. coli* reveal regulatory architectures governed by metabolism | * The paper never explains what are the populations in their data, ie what Evo04pgiEvo04EP stands for. * supplementary data 6: 918 mutations listed (but this is supposed to be a file with 'mutation statistics' which it is not - note that the names of the populations here don't match to the paper or support data 5) * the number of mutations listed in the paper does not match any number of mutations given in the SI |
| Sandberg 2016 | Evolution of *E*. *coli* on [U-13C] Glucose Reveals a Negligible Isotopic Influence on Metabolism and Physiology | * The paper never mentions the amount of mutations that occur, all that I have to work on is an average of 6 mutations per strain. |
| Sherlock 2013 | Whole Genome, Whole Population Sequencing Reveals That Loss of Signaling Networks Is the Major Adaptive Strategy in a Constant Environment | * Paper refers to 120 mutations which doesn’t match the number of mutations is the dataset. * Then the paper says it has 74 of the identified mutations (63%) decrease in frequency following their maxima, and 42 of these mutations (57%) become extinct by the end of the experiment (74+42 = 116), which still doesn’t add up to 120. * The dataset has 117 mutations, which doesn’t add up with any of the information in the paper. |
| Sherlock 2019 | Evolutionary dynamics of de novo mutations and mutant lineages arising in a simple, constant environment | * Paper indicates there should be 3346 mutations, the dataset has only 3327 mutations. |
| Tonoyan 2019 | Continuous culture of *Escherichia coli*, under selective pressure by a novel antimicrobial complex, does not result in development of resistance | * Paper indicates there should be a certain number of mutations per population but it does not match the number of mutations in the dataset. * Only mutation information provided for the thawed cultures. |
| LaCroix 2014 | Use of Adaptive Laboratory Evolution To Discover Key Mutations Enabling Rapid Growth of *Escherichia coli* K-12 MG1655 on Glucose Minimal Medium | * The data provided for the number of mutations has several different flasks in the same population. Nowhere in the paper it indicates why they chose to pick these flasks. From population to population, the number on the flasks as well as the number of flasks are different. |
| Herron 2013 | Parallel Evolutionary Dynamics of Adaptive Diversification in *Escherichia coli* | * They didn't break down the information by generation, just compiled all into one column of a dataset. So there is no way without using data thief to determine presence/absence across the time series of 16 time points. * Missing information about the populations. |