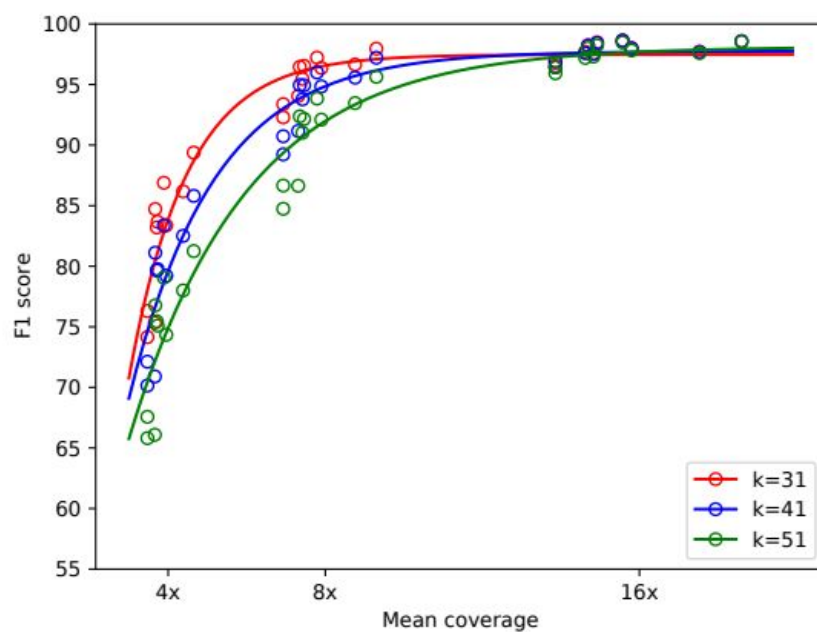
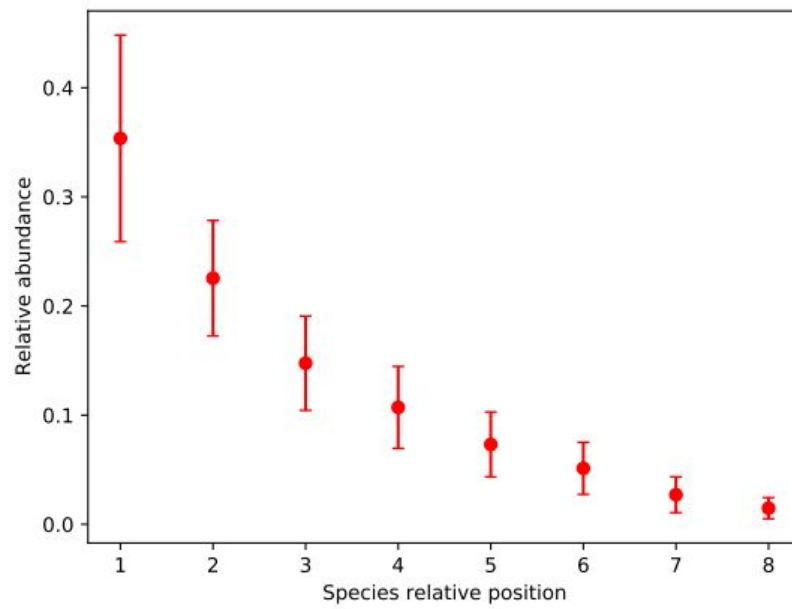


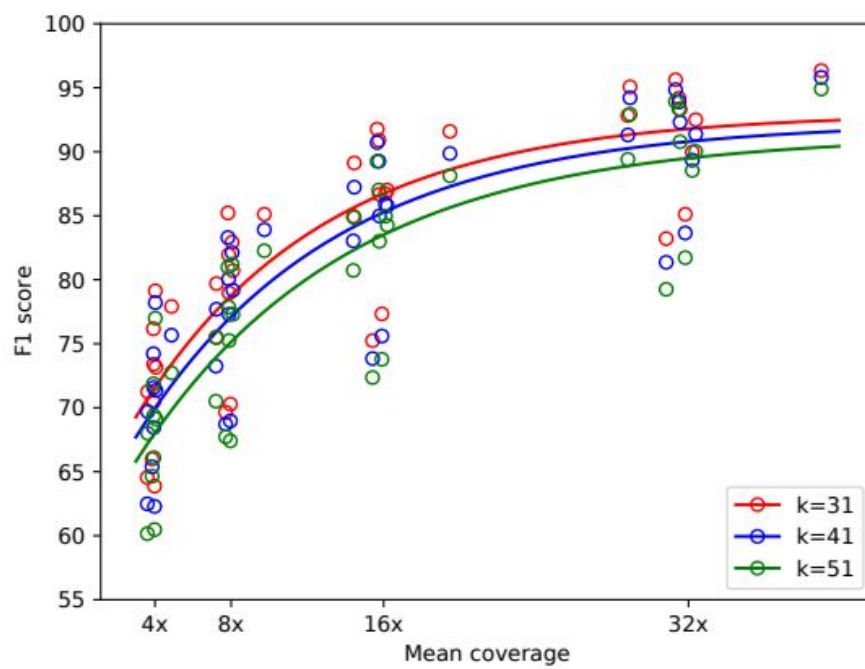
**Figure S1. Dependence of bases quality phred score on the position in read for simulated datasets.**



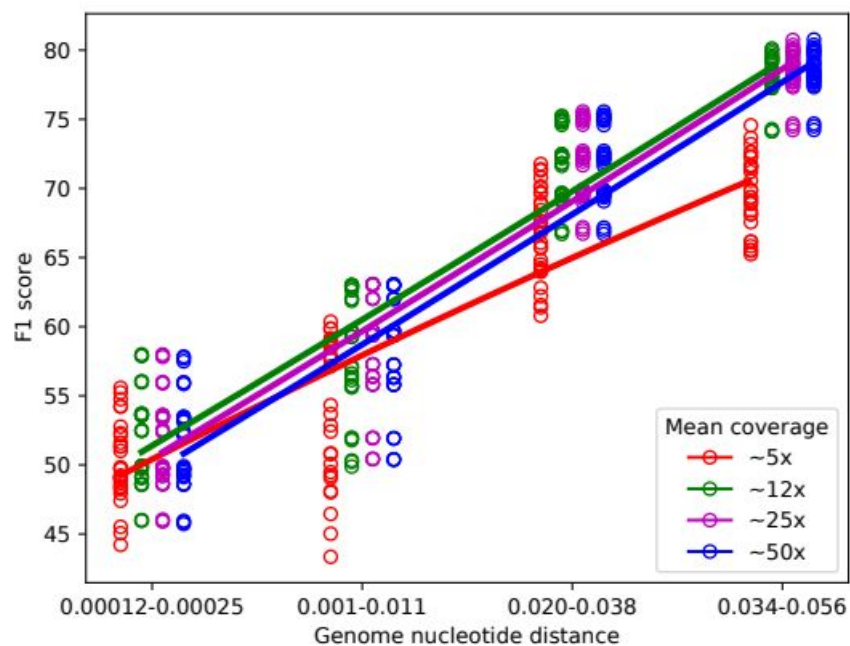
**Figure S2. Accuracy on simulation dataset from species bacterial genomes with uniform relative abundance distribution.**



**Figure S3. Relative abundance of species drawn from exponential distribution.**



**Figure S4. Accuracy on simulation dataset from species with exponential relative abundance distribution.**



**Figure S5. Accuracy on simulation dataset from strains with different Mash distances.**