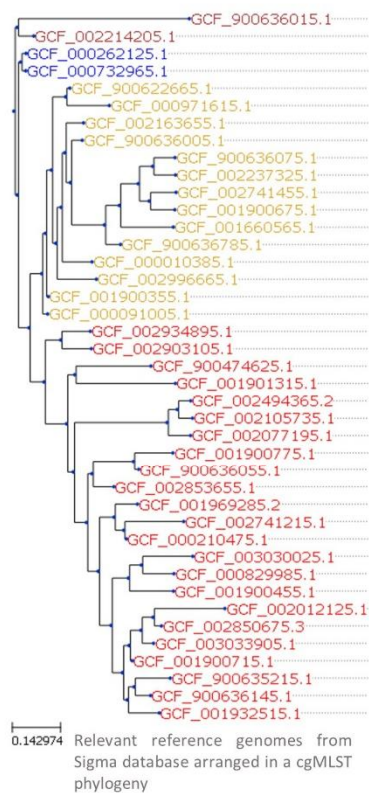


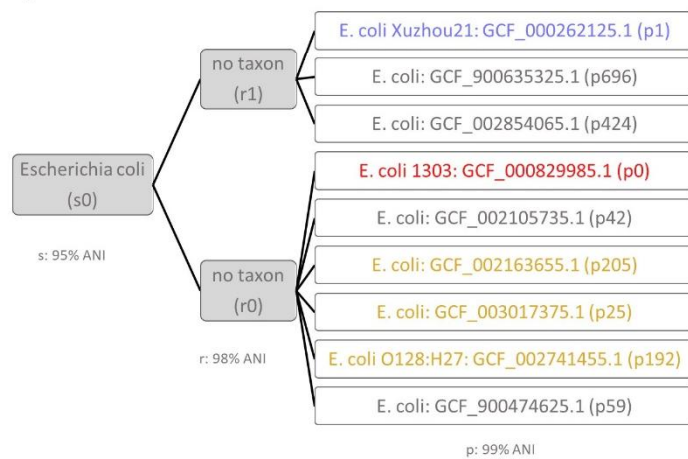
Sigma



Reference Genome	spMm24h	Mm24h
GCF_900636015.1	562 (0.0%, 0.1%)	668 (0.0%, 0.2%)
GCF_002214205.1	0 (0.0%, 0.0%)	946 (0.0%, 0.2%)
GCF_000262125.1	2416 (0.1%, 0.5%)	0 (0.0%, 0.0%)
GCF_000732965.1	269278 (6.3%, 52.8%)	0 (0.0%, 0.0%)
GCF_900622665.1	358 (0.0%, 0.1%)	0 (0.0%, 0.0%)
GCF_000971615.1	312 (0.0%, 0.1%)	0 (0.0%, 0.0%)
GCF_002163655.1	680 (0.0%, 0.1%)	0 (0.0%, 0.0%)
GCF_900636005.1	454 (0.0%, 0.1%)	0 (0.0%, 0.0%)
GCF_900636075.1	9522 (0.2%, 1.9%)	0 (0.0%, 0.0%)
GCF_002237325.1	638 (0.0%, 0.1%)	0 (0.0%, 0.0%)
GCF_002741455.1	163200 (3.8%, 32.0%)	0 (0.0%, 0.0%)
GCF_001900675.1	1830 (0.0%, 0.4%)	0 (0.0%, 0.0%)
GCF_001660565.1	1242 (0.0%, 0.2%)	0 (0.0%, 0.0%)
GCF_900636785.1	2286 (0.1%, 0.4%)	0 (0.0%, 0.0%)
GCF_000010385.1	464 (0.0%, 0.1%)	0 (0.0%, 0.0%)
GCF_002996665.1	994 (0.0%, 0.2%)	0 (0.0%, 0.0%)
GCF_001900355.1	1586 (0.0%, 0.3%)	0 (0.0%, 0.0%)
GCF_000091005.1	728 (0.0%, 0.1%)	0 (0.0%, 0.0%)
GCF_002934895.1	1150 (0.0%, 0.2%)	0 (0.0%, 0.0%)
GCF_002903105.1	0 (0.0%, 0.0%)	1600 (0.0%, 0.4%)
GCF_900474625.1	1290 (0.0%, 0.3%)	0 (0.0%, 0.0%)
GCF_001901315.1	0 (0.0%, 0.0%)	3112 (0.1%, 0.8%)
GCF_002494365.2	0 (0.0%, 0.0%)	1206 (0.0%, 0.3%)
GCF_002105735.1	4684 (0.1%, 0.9%)	12682 (0.3%, 3.1%)
GCF_002077195.1	0 (0.0%, 0.0%)	296 (0.0%, 0.1%)
GCF_001900775.1	0 (0.0%, 0.0%)	458 (0.0%, 0.1%)
GCF_900636055.1	0 (0.0%, 0.0%)	1284 (0.0%, 0.3%)
GCF_002853655.1	0 (0.0%, 0.0%)	926 (0.0%, 0.2%)
GCF_001969285.2	516 (0.0%, 0.1%)	284 (0.0%, 0.1%)
GCF_002741215.1	314 (0.0%, 0.1%)	0 (0.0%, 0.0%)
GCF_000210475.1	272 (0.0%, 0.1%)	0 (0.0%, 0.0%)
GCF_003030025.1	0 (0.0%, 0.0%)	214 (0.0%, 0.1%)
GCF_000829985.1	37162 (0.9%, 7.3%)	380148 (7.7%, 91.6%)
GCF_001900455.1	0 (0.0%, 0.0%)	744 (0.0%, 0.2%)
GCF_002012125.1	0 (0.0%, 0.0%)	318 (0.0%, 0.1%)
GCF_002850675.3	0 (0.0%, 0.0%)	450 (0.0%, 0.1%)
GCF_003033905.1	1522 (0.0%, 0.3%)	0 (0.0%, 0.0%)
GCF_001900715.1	768 (0.0%, 0.2%)	7154 (0.1%, 1.7%)
GCF_900635215.1	302 (0.0%, 0.1%)	632 (0.0%, 0.2%)
GCF_900636145.1	4148 (0.1%, 0.8%)	0 (0.0%, 0.0%)
GCF_001932515.1	0 (0.0%, 0.0%)	1646 (0.0%, 0.4%)

Cluster	spMm24h	Mm24h	
spMm24h_Sigma_cl4	562 (0.01%, 0.1%)	Mm24h_Sigma_cl4	1614 (0.03%, 0.4%)
spMm24h_Sigma_cl2	271694 (6.4%, 53.3%)		
spMm24h_Sigma_cl3	184294 (4.3%, 36.1%)		
spMm24h_Sigma_cl1	52128 (1.2%, 10.2%)	Mm24h_Sigma_cl1	413154 (8.39%, 99.6%)

Sparse



Selected pre-defined hierarchically organized clusters used in Sparse reference database

Cluster	spMm24h	Mm24h	
spMm24h_Sparse_p1	230990 (5.43%, 53.94%)		
spMm24h_Sparse_p696	1748 (0.04%, 0.41%)		
spMm24h_Sparse_p424	1658 (0.04%, 0.39%)		
spMm24h_Sparse_p0	78588 (1.85%, 18.35%)	Mm24h_Sparse_p0	327746 (6.66%, 99.34%)
		Mm24h_Sparse_p42	2162 (0.04%, 0.66%)
spMm24h_Sparse_p205	15156 (0.36%, 3.54%)		
spMm24h_Sparse_p25	15148 (0.36%, 3.54%)		
spMm24h_Sparse_p192	79812 (1.88%, 18.64%)		
spMm24h_Sparse_p59	2918 (0.07%, 0.68%)		

Extracted clusters, and corresponding read numbers (% of total reads, % of reads assigned to *E. coli* reference genomes)

Supplementary Figure 2. Strain-level metagenomic analysis of enriched minced meat sample (Mm24h) and an enriched minced meat sample that has been spiked with a pathogenic *E. coli* isolate TIAC1152 (spMm24h) using Sigma and Sparse. Above: number of reads from Mm24h and spMm24h samples assigned by Sigma to each reference genome from the database containing 728 complete *E. coli* genomes. A cgMLST phylogeny of the reference genomes to which at least 0.05% of *E. coli* reads were assigned for any of the two processed samples is shown on the left. In the middle, the number of Mm24h and spMm24h reads assigned to each reference genome are listed. Reads from groups of

closely related reference genomes were pooled to generate the clusters listed on the right. The colors of the references and of the clusters correspond to the colors used in Supplementary Figure 1, allowing to identify the section of the reference genome cgMLST tree from which the underlying references originated. Below: number of reads from Mm24h and spMm24h assigned by Sparse to each cluster from the database of 728 hierarchically ordered complete *E. coli* genomes. The clusters of assemblies with an average nucleotide identity (ANI) of 95%, 98% and 99% to which any *E. coli* reads were assigned for any of the two processed samples are shown on the left. Resulting clusters are shown on the right. The Sparse clusters containing a considerable amount of reads carry the same color as the corresponding Sigma clusters (see also Figure 1).