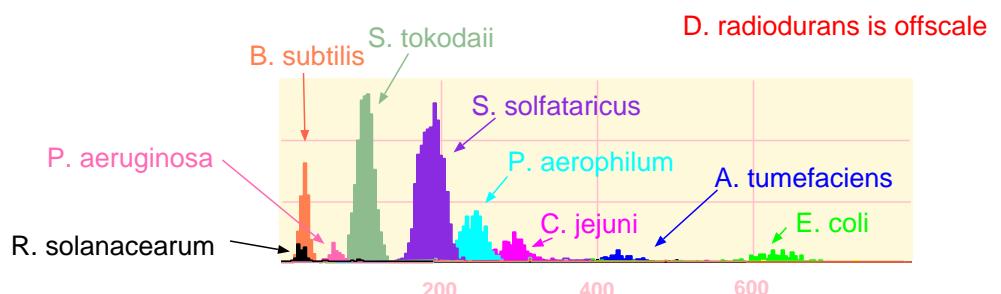


Tag frequency histograms of oversampled GSTs from a mixture of genomes yield peaks of tags, one peak per organism. The means of the peaks are simply the relative abundances of the organisms in the mix. The total area under each peak is a product of the number of GSTs in the organisms genome and the number of samples.

This series of histograms for Monte Carlo simulation of drawing GSTs (Spel fragment, Nlalll anchor) demonstrates this effect. The organisms and abundances used are:

800K samples is sufficient to isolate 9 out 10 organisms and determine abundances



Organism	# GSTs	Abundance
<i>Deinococcus radiodurans</i>	42	41.1%
<i>Escherichia coli</i> K12	140	18.2
<i>Agrobacterium tumefaciens</i> C58	96	12.7
<i>Campylobacter jejuni</i>	262	8.6
<i>Pyrobaculum aerophilum</i>	452	7.1
<i>Pseudomonas aeruginosa</i>	74	2.0
<i>Sulfolobus solfataricus</i>	1518	5.6
<i>Sulfolobus tokodaii</i>	1156	3.1
<i>Bacillus subtilis</i>	308	0.9
<i>Ralstonia solanacearum</i>	72	0.8