

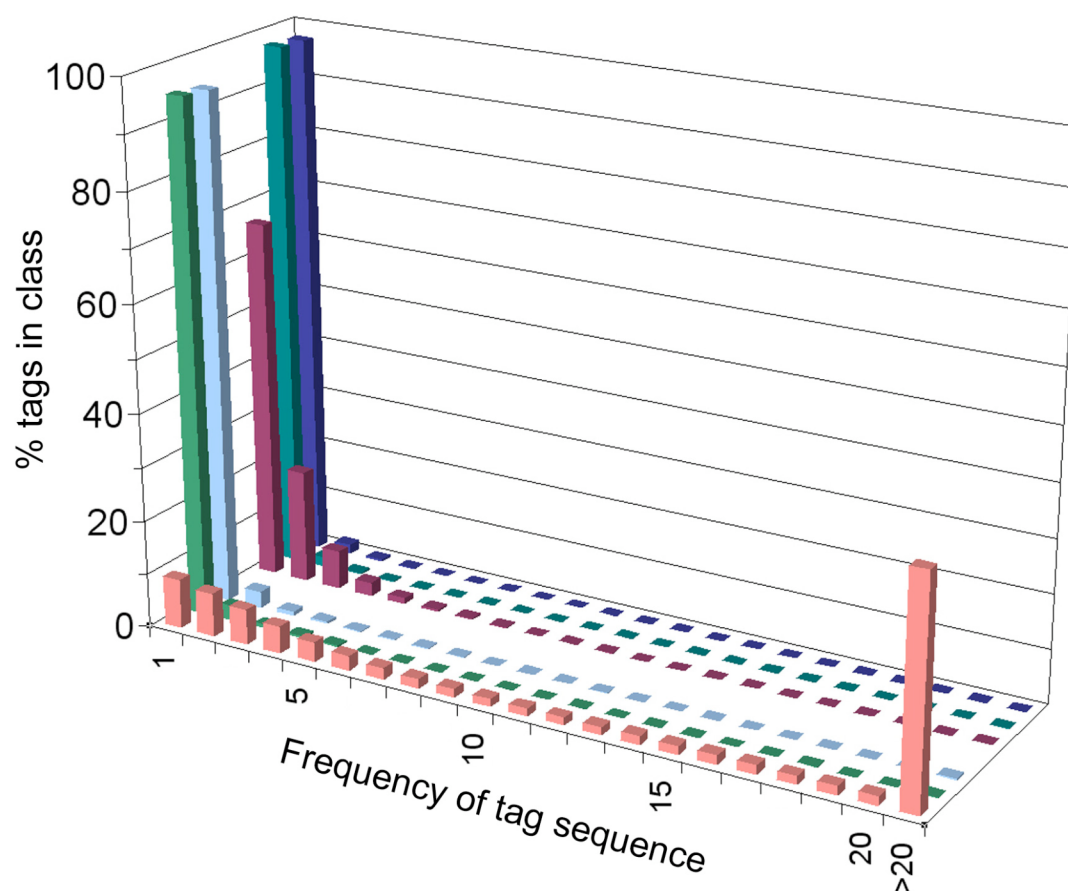


Figure S1. *Effect of tag length on frequency of matches to the genome and transcriptome.*

All possible tags were extracted from the Ensembl genome and transcriptome. The frequency of each tag in the genome or transcriptome was calculated and tags were grouped by frequency. The different tag lengths chosen correspond to the different techniques used in this thesis: 14 bases, SAGE; 21 bases, LongSAGE; and 20 bases, MPSS.

	Tags Extracted From	Restriction Site	Tag Length (bases)
	Genome	CATG	14
	Genome	CATG	21
	Genome	GATC	20
	Transcriptome	CATG	14
	Transcriptome	CATG	21
	Transcriptome	GATC	20