Table S3 Comparison of gene-level abundance results from alignment of 30 million simulated CAST reads to NCBIM37 and CAST transcriptomes

			Number of genes with estimates x% from Ground Truth			
Aligned to	**Mismatches Allowed**	**Genes above threshold**	**< 5%**	**< 10%**	**> 10%**	**> 50%**
30M CAST Reads						
NCBIM37	3	13,848	3,701	7,850 (57%)	5,998 (43%)	654
CAST	3	13,756	10,040	11,939 (87%)	1,794 (13%)	272
NCBIM37	0	13,788	1,535	3,127 (23%)	10,661 (77%)	2,082
CAST	0	13,738	9,322	11,325 (82%)	2,386 (18%)	259

Alignment of 30 million simulated CAST reads to the individualized CAST transcriptome (≤3 mismatches) results in nearly three times as many gene estimates (N= +6,339) that fall within 5% of ground truth value and fewer than a third as many gene estimates (N= -4,204) that deviate more than 10% from the ground truth. Gene-level abundance results for perfect matching reads (i.e. 0 mismatches) are also shown.