Table S11 Isoform abundance results in CAST simulation study

10 Million Simulated CAST reads

			Number of isoforms with estimates x% from Ground Truth			
Aligned to	**Mismatches Allowed**	**Isoforms above threshold**	**< 5%**	**< 10%**	**> 10%**	**> 50%**
NCBIM37	3	21,568	3,908	6,581 (30%)	14,987 (70%)	7,096
CAST	3	21,457	3,244	7,796 (36%)	13,661 (64%)	6,551
NCBIM37	0	21,363	1,393	2,883 (13%)	18,480 (87%)	9,488
CAST	0	21,222	1,998	5,089 (24%)	16,133 (76%)	6,540

30 Million Simulated CAST reads

			Number of isoforms with estimates x% from Ground Truth			
Aligned to	**Mismatches Allowed**	**Isoforms above threshold**	**< 5%**	**< 10%**	**> 10%**	**> 50%**
NCBIM37	3	27,048	3,600	7,217 (27%)	19,831 (73%)	9,821
CAST	3	26,910	6,685	9,951 (37%)	16,959 (63%)	9,031
NCBIM37	0	26,909	1,765	3,454 (13%)	23,455 (87%)	12,748
CAST	0	26,695	6,792	9,578 (36%)	17,013 (64%)	8,821

Alignment of simulated CAST reads to the individualized CAST transcriptome (≤3 mismatches) improves estimates of isoform abundance compared to alignment to NCBIM37. Increasing the sequencing depth from 10 to 30 million single-end reads significantly does not improve isoform resolution – more isoform estimates fall within five percent of the simulated ground truth but the total number of isoforms expressed above threshold increases too, causing no relative improvement in the accuracy of isoform abundance estimates. Isoform-level abundance results for perfect matching reads (i.e. 0 mismatches) are also shown.