Table S1: **Percent of Exons Across the Possible Values for the Number of Non-boundary Samples in a 5 versus 5 setting.** Maximum value for each exon is 10 corresponding to no samples lying on the boundaries. Minimum value is 0 corresponding to an exon where all samples lie on the boundaries $\{0, 1\}$. n_{eff} is calculated as the maximum of the number of non-boundary samples and $K + 1$, where K is the number of groups. For this paper K=2 corresponding to a minimum value of 3 for n_{eff}, given to exons with only 0, 1 or 2 non-boundary samples.

Data Set	Possible Values for Number of Non-boundary Samples (5 versus 5)										
Type	0	1	2	3	4	5	6	7	8	9	10
Double Bin. Simulation	0.02	10.81	9.51	8.43	7.70	7.45	7.10	7.47	7.97	10.43	23.12
Beta Bin. Simulation	0.00[†]	15.92	7.36	4.65	3.63	3.24	3.17	3.31	4.53	9.07	45.12
Tissue Data	1.13	10.95	8.60	7.78	7.50	7.44	7.19	7.62	8.67	10.58	22.52

[†]Percentage is exactly zero.

Table S2: **Comparison of Double Binomial-based Methods.** Shown in the table below are the average percentage of exons called significant after multiple testing adjustments from the Tissue Data under 100 simulations of null and real scenarios described above for the methods we developed based on the double binomial distribution. The total number of exons is $412,002$. The rates are percentages out of only those exons that had at least one skipping event, a number which varies with sample size but is roughly 1/4 of all exons. See Supplemental Table S3 for both the precise number of exons called and analyzed.

Sample	WEB-Seq		DEB-Seq		Wt-Likelihood	
Size	Real	Null	Real	Null	Real	Null
2 vs 2	0.89	0.00	1.57	0.00	1.93	0.03
3 vs 3	2.82	0.00	4.55	0.00	4.06	0.02
4 vs 4	6.73	0.00	8.86	0.00	7.13	0.00
5 vs 5	12.16	0.00	13.78	0.00	10.85	0.00
6 vs 6	15.94	0.00	17.28	0.00	13.64	0.00
7 vs 7	20.26	0.00	21.24	0.00	16.79	0.00

Table S3: **Comparison of the Total Calls for Double Binomial Based Methods.** Shown are the explicit number of significant calls analyzing the Tissue Data under the Null and Real Scenarios. The percentages in Table S2 in the main text are out of only those exons that had at least one skipping event, a number which varies with sample size and is given in this table.

Sample	WEB-Seq		DEB-Seq		Wt-Likelihood		Total #
Size	Real	Null	Real	Null	Real	Null	of Exons
2 vs 2	1,112	6	1,972	0	2,426	40	125,398
3 vs 3	3,535	4	5,708	0	5,094	19	125,398
4 vs 4	8,445	0	11,105	0	8,940	1	125,398
5 vs 5	15,252	1	17,281	0	13,612	1	125,398
6 vs 6	19,984	1	21,663	1	17,107	5	125,398
7 vs 7	25,401	1	26,635	1	21,058	1	125,398

Table S4: **Comparison of the Total Calls for Competing Methods.** Shown are the explicit number of significant calls analyzing the Tissue Data under the Null and Real Scenarios for the competing methods. The top table gives the results for all methods that take as input the same set of inclusion-exclusion counts as in Table S3, including DEXSeq after post-filtering to the same set of exons. The bottom table gives the results for the two methods which have different numbers of exons than the comparisons in Table S3 (the unfiltered DEXSeq and MATS), and their total exon numbers are indicated in the table.

Sample	DEXSeq*		DSS		EB2		BBSeq		Quasi-bin.		Beta-bin.	
Size	Real	Null	Real	Null	Real	Null	Real	Null	Real	Null	Real	Null
2 vs 2	16,931	1,825	13,075	1,643	13,892	9,365	23,545	8,232	425	168	6,346	0
3 vs 3	21,525	775	15,286	732	14,492	9,032	31,284	9,236	1,761	94	12,162	6
4 vs 4	27,302	148	18,180	235	14,974	8,589	36,243	7,773	9,262	70	17,218	0
5 vs 5	33,392	99	21,787	200	15,883	8,198	39,222	7,555	19,130	72	22,232	0
6 vs 6	37,618	141	24,372	221	16,643	7,850	40,911	7,188	25,466	93	25,593	10
7 vs 7	42,460	81	27,631	173	17,658	7,523	42,408	6,347	31,631	123	29,370	1

*Post Filtered to have same set of exons as inclusion/exclusion setting.

Sample	DEXSeq			MATS**	
Size	Real	Null	Total #	Null	Total #
2 vs 2	74,754	6,505	412,002	1,168	33,854
3 vs 3	90,885	2,521	412,002	646	36,396
4 vs 4	111,966	460	412,002	936	38,170
5 vs 5	131,868	292	412,002	1,070	38,987
6 vs 6	145,195	346	412,002	1,557	39,590
7 vs 7	159,216	85	412,002	1,363	40,190

**WEB-Seq makes at most one significant call on this set of exons (for sample sizes 3, 5 & 7) and zero for other sample sizes.

Table S5: **Percent of Boundary Calls Out of Total Significant.** Shown below are the percentage of significant calls for which the exon has a log-Fold-Change of the odds between the groups that is infinite, corresponding to a situation where all samples of at least one of the two groups lie on the boundary, either all 1 or all 0. For DEXSeq, the percentage is based on the significance results after post-filtering to the DEXSeq results to the same set of exons as the inclusions/exclusion counts, i.e. those exons with non-zero skipping in some sample. For 2 vs 2, WEB-Seq has no significant calls, so the percentage is not defined.

Sample Size	DEXSeq	EB2	BBSeq	Quasi-bin.	Beta-bin.	WEB-Seq
2 vs 2	0.43	12.42	0.18	11.70	0.00[†]	-
3 vs 3	0.78	12.78	0.55	14.68	1.21	0.23
4 vs 4	0.87	12.64	0.71	4.58	1.61	0.42
5 vs 5	1.74	18.33	0.00[†]	9.38	3.36	0.76
6 vs 6	1.77	11.16	1.11	6.85	2.87	1.25
7 vs 7	1.39	7.57	1.66	4.66	2.24	1.27
10 vs 10	1.91	4.55	1.93	5.14	2.77	2.28
15 vs 15	3.06	1.88	2.03	5.64	2.96	3.92

[†]Percentage is exactly zero.

Table S6: **Percent of single-exon calls made by DEXSeq, by annotation and skipping event.**

	% of Total Single-Exon Calls				% of Total Significant Calls			
Sample Size	2 vs 2	3 vs 3	5 vs 5	15 vs 15	2 vs 2	3 vs 3	5 vs 5	15 vs 15
Alternatively Spliced	81.75	70.74	67.70	54.26	85.89	87.20	87.01	88.11
Non-skipped Constitutive	18.00	29.06	32.12	45.74	13.88	12.58	12.79	11.72
Skipped Constitutive	0.25	0.19	0.18	0.00[†]	0.23	0.22	0.20	0.17
Total Calls	2,356	2,085	1,681	916	16,229	59,871	87,144	186,570

[†]Percentage is exactly zero.