	Pr(r reads from UMI will produce >=2 UEIs of >=2 reads each)								Mean error UMI-variants			
	SSD of same								Mean reads per UMI-variant			
	<= 3 reads	4 reads	5 reads	6 reads	7 reads	8 reads	9 reads	10 reads	1 errors	2 errors	3 errors	**FPR**
Sample 1	0.0E+00	1.6E-01	3.5E-01	5.3E-01	6.6E-01	7.5E-01	8.2E-01	8.6E-01	9.4E-01	1.4E-02	4.8E-05	**9.5E-05**
	0.0E+00	4.2E-04	4.9E-04	4.6E-04	4.1E-04	2.9E-04	4.6E-04	2.9E-04	1.7E+00	1.2E+00	1.0E+00	
Sample 2	0.0E+00	1.7E-01	3.7E-01	5.5E-01	6.8E-01	7.7E-01	8.4E-01	8.8E-01	9.4E-01	1.3E-02	3.1E-05	**9.3E-05**
	0.0E+00	6.3E-04	5.7E-04	3.5E-04	3.9E-04	4.0E-04	3.7E-04	4.3E-04	1.7E+00	1.2E+00	1.0E+00	
Sample 3	0.0E+00	1.1E-01	2.5E-01	3.9E-01	5.0E-01	6.0E-01	6.7E-01	7.3E-01	6.3E+00	3.6E-01	3.5E-03	**1.3E-03**
	0.0E+00	1.1E-03	1.7E-03	1.2E-03	2.3E-03	1.8E-03	9.2E-04	1.1E-03	3.2E+00	1.1E+00	1.0E+00	
Sample 4	0.0E+00	1.5E-01	3.2E-01	4.7E-01	5.9E-01	6.7E-01	7.4E-01	7.8E-01	3.1E+00	6.2E-02	1.8E-04	**3.3E-04**
	0.0E+00	7.4E-04	7.8E-04	1.0E-03	7.0E-04	9.9E-04	8.0E-04	9.8E-04	2.0E+00	1.1E+00	1.0E+00	
Sample 5	0.0E+00	1.5E-01	3.3E-01	4.8E-01	5.9E-01	6.7E-01	7.4E-01	7.8E-01	7.5E+00	2.4E-01	1.6E-03	**1.5E-03**
	0.0E+00	8.8E-04	1.9E-03	1.8E-03	1.2E-03	1.2E-03	1.6E-03	9.2E-04	2.5E+00	1.2E+00	1.1E+00	

TABLE S4: False-positive rate estimates for each sample using equation 3, related to Figures 2, 4-6, and 5. "Mean error UMI-variants" corresponds to quantity μ_k, with k being the number of errors, and "mean reads per UMI-variant" corresponds to quantity $\mu_{R,k}$. FPR is then the estimated fraction of UMIs that are mistakenly identified as real in the final data set.