

Table S2 Summary by coding versus noncoding regions for each coverage class.

Coverage Class ^a	Count	<i>L</i> total	<i>L</i> coding	<i>L</i> noncoding	Masked <i>L</i> coding ^b	Masked <i>L</i> noncoding ^b	<i>S</i> coding	<i>S</i> noncoding	Masked <i>S</i> coding ^b	Masked <i>S</i> noncoding ^b
18	541	183938	125940	57998	138	76	1045	641	0	0
17	422	153028	100226	52802	262	131	891	674	2	1
16	302	111593	71956	39637	220	74	521	568	0	0
15	218	84201	56940	27261	232	62	474	344	2	0
14	192	74973	45635	29338	293	154	360	498	0	1
13	140	54967	38671	16296	205	30	361	211	1	0
12	160	63364	42051	21313	364	142	353	229	3	0
11	120	48291	30583	17708	334	115	247	222	1	0
10	115	44879	30935	13944	491	54	227	137	1	0
9	88	38335	24047	14288	163	71	254	139	1	0
8	103	41122	23613	17509	132	161	165	158	1	0
7	79	31723	21421	10302	129	101	257	134	1	0
6	80	34431	21666	12765	262	136	283	140	1	1
5	103	42626	27736	14890	254	198	222	169	3	4
4	93	37578	24879	12699	193	89	198	94	0	0
3	95	41177	27362	13815	223	159	208	86	1	0
2	136	56143	36340	19803	277	93	206	95	0	0
1	174	75124	46686	28438	197	139	NA	NA	NA	NA

Abbreviations: *L*, length; *S*, segregating sites.

^aSample size in the alignment (i.e. the number of sequences).

^bMasked bases are the number of aligned sites with at least one base masked due to its quality score < 30.