Supplementary Materials

Supplementary Files

Supplementary File S1. List of mutations. A zip file containing four CSV files: a list of unique, shared mutations, the mutations that found in all 107 MA lines, and the overlap of MA line SNMs with variants in the 1001 Genomes populations.

Supplementary File S2. A single VCF file for variants in the 107 MA lines.

Supplementary File S3. VCF files for variants and non-variants in each individual MA lines, compressed in 14 tar files.

Supplementary File S4. A single VCF file for variants in five sub-founder lines.

Supplementary File S5. Simulated mutations. A zip file containing five reference genomes, each with 100 simulated point mutations, and five VCF files for variants called in each of the simulated reference genomes.

Supplementary File S6. Bash and R scripts.

Supplementary Figures

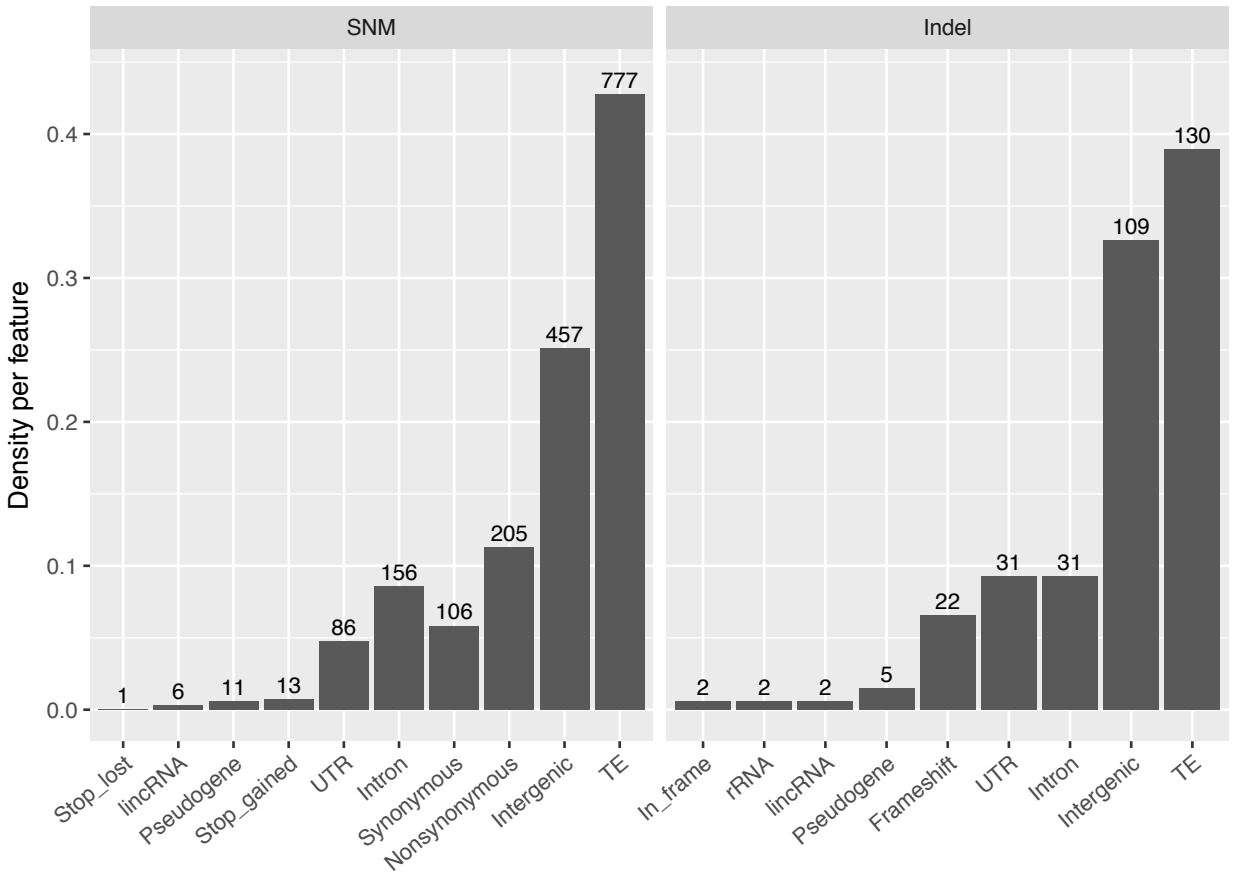

Figure S1. Distribution of indels and SNMs from 107 MA lines in different genomic regions. Numbers on top of bars indicates absolute numbers of SNMs or indels. lincRNA: long intergenic non coding RNA. TE: transposable elements. UTR: 3' and 5' untranslated regions.

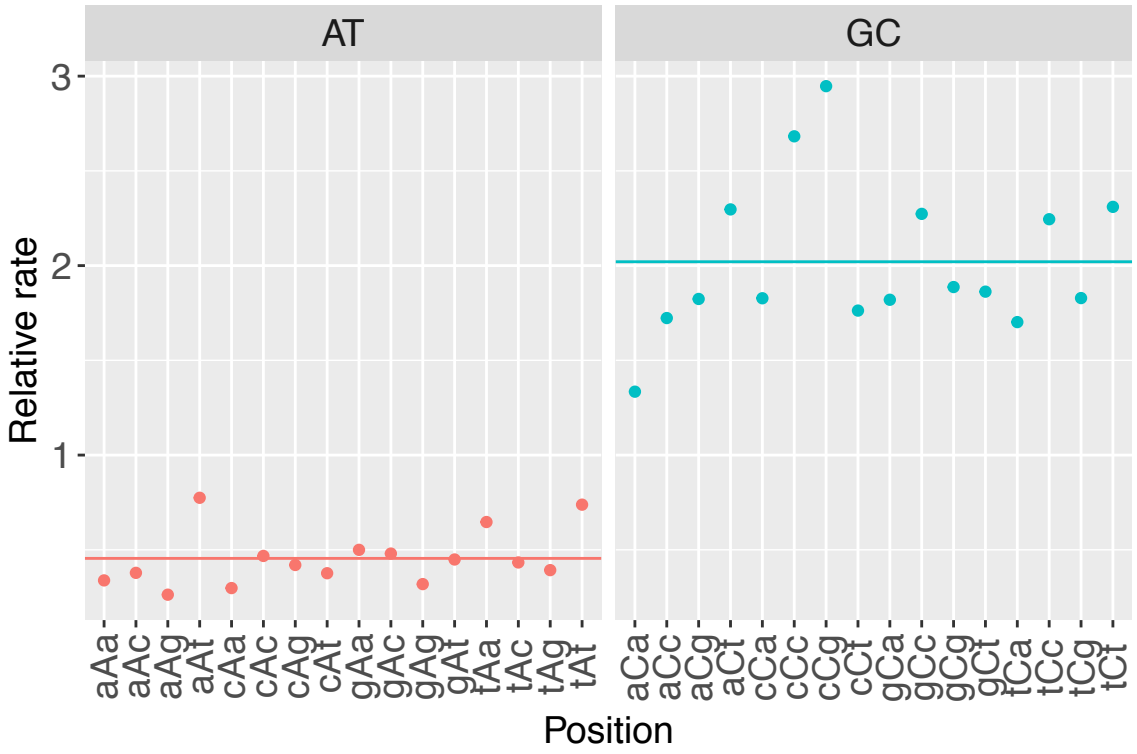

Figure S2. The trinucleotide-context dependent mutation rate at AT and GC bases estimated from 107 MA lines. The mutation rate is shown relative to the genome-wide SNM mutation rate. The x-axis shows the focal bases (upper case) and the immediately flanking bases (lower case), regardless of the strand orientation (e.g., aAa class includes mutation rates at aAa and tTt classes). The horizontal lines show the average mutation rates for AT (red) and GC (green) classes. Regardless of the flanking nucleotides, GC bases have significantly higher mutation rates than AT bases (two sample Student's t-test, t = 14.4, df = 18.9, p = 1e-11).

Supplementary Tables

Supplementary Table S1. Estimated mutation rates for each MA line.

MA Line	[a]Total indel	[a]Total SNM	Read depth coverage mean	Percentage of bases > 15x coverage	Accessible reference sites	Indel mutation rate	SNM mutation rate
1	2 (1)	17 (12)	42.38	99.1	113846082	7.03E-10	5.97E-09
2	6 (3)	14 (12)	50.18	99.6	114405538	2.10E-09	4.89E-09
3	6 (1)	18 (4)	41.42	98.7	114080507	2.10E-09	6.31E-09
4	1 (1)	22 (1)	42.67	98.7	114090170	3.51E-10	7.71E-09
5	3 (0)	32 (6)	51.9	99.7	114686793	1.05E-09	1.12E-08
6	3 (0)	24 (6)	39.04	97	113966211	1.05E-09	8.42E-09
7*	6 (0)	21 (0)	48.58	99.6	114652290	2.09E-09	7.33E-09
8*	4 (0)	24 (0)	46.8	99.6	114418276	1.40E-09	8.39E-09
9	0 (0)	27 (20)	53.77	99.8	114671951	0.00E+00	9.42E-09
11	2 (1)	14 (12)	21.28	41.1	110787995	7.22E-10	5.05E-09
13*	4 (0)	20 (0)	21.44	42.8	111935961	1.43E-09	7.15E-09
14	3 (0)	16 (1)	78.09	99.9	113483773	1.06E-09	5.64E-09
15	4 (4)	16 (9)	30.73	88.3	113275511	1.41E-09	5.65E-09
16*	1 (0)	19 (0)	15.55	7.8	103907260	3.85E-10	7.31E-09
17	7 (4)	20 (9)	27.73	77	112840675	2.48E-09	7.09E-09
18*	1 (0)	23 (0)	24.16	50.9	112424522	3.56E-10	8.18E-09
19*	4 (0)	25 (0)	30.38	82.1	112468981	1.42E-09	8.89E-09
20	8 (0)	32 (1)	37.53	95.7	113312408	2.82E-09	1.13E-08
22	2 (1)	17 (8)	24.49	50.1	111360652	7.18E-10	6.11E-09
23	3 (1)	19 (8)	25.83	66.5	112465906	1.07E-09	6.76E-09
24*	1 (0)	26 (0)	28.65	76.7	112793272	3.55E-10	9.22E-09
25	5 (1)	20 (3)	36.67	95.2	113277738	1.77E-09	7.06E-09
26	4 (0)	17 (1)	26.53	63.9	111270274	1.44E-09	6.11E-09
27*	6 (0)	16 (0)	26.31	57.8	111486451	2.15E-09	5.74E-09
28	4 (0)	9 (1)	15.99	5.5	99934668	1.60E-09	3.60E-09
29	4 (0)	18 (3)	16.22	11	104627523	1.53E-09	6.88E-09
31*	1 (0)	13 (0)	17.44	17.5	97366818	4.11E-10	5.34E-09
32*	11 (0)	40 (0)	18.75	20.3	107597083	4.09E-09	1.49E-08
33*	4 (0)	16 (0)	18.45	20.4	108756412	1.47E-09	5.88E-09
34*	3 (0)	20 (0)	34.36	95	113580101	1.06E-09	7.04E-09
35*	2 (0)	19 (0)	37.45	96.2	111027064	7.21E-10	6.85E-09
36*	5 (0)	14 (0)	14.9	6.9	72218737	2.77E-09	7.75E-09
37*	3 (0)	10 (0)	14.53	6.1	100432149	1.19E-09	3.98E-09
38*	5 (0)	18 (0)	15.01	7.1	101475323	1.97E-09	7.10E-09
39	1 (1)	12 (8)	23.99	53.6	93728864	4.27E-10	5.12E-09

MA Line	[a]Total indel	[a]Total SNM	Read depth coverage mean	Percentage of bases > 15x coverage	Accessible reference sites	Indel mutation rate	SNM mutation rate
40*	3 (0)	15 (0)	61.96	99.8	114188113	1.05E-09	5.25E-09
41*	1 (0)	12 (0)	71.57	99.9	114999211	3.48E-10	4.17E-09
42*	4 (0)	16 (0)	66.97	99.9	114599852	1.40E-09	5.58E-09
43*	4 (0)	29 (0)	44.84	99.1	113630269	1.41E-09	1.02E-08
44	8 (1)	10 (4)	42.7	97.6	113243513	2.83E-09	3.53E-09
45*	3 (0)	19 (0)	41.87	97.3	113105551	1.06E-09	6.72E-09
46	0 (0)	15 (1)	38.31	96	114090044	0.00E+00	5.26E-09
47	3 (1)	26 (4)	37.31	95	114022308	1.05E-09	9.12E-09
48*	1 (0)	17 (0)	64.25	99.9	114906640	3.48E-10	5.92E-09
49	5 (0)	16 (1)	58.83	99.8	114697489	1.74E-09	5.58E-09
50*	3 (0)	16 (0)	67.79	99.9	114640888	1.05E-09	5.58E-09
51	1 (0)	13 (3)	34.53	93.2	113886861	3.51E-10	4.57E-09
52*	4 (0)	19 (0)	22.57	35.1	111460176	1.44E-09	6.82E-09
53	2 (0)	25 (20)	38.6	97.1	113984996	7.02E-10	8.77E-09
54	4 (0)	19 (3)	32.73	84.7	113431315	1.41E-09	6.70E-09
55*	2 (0)	19 (0)	35.16	94.5	113894515	7.02E-10	6.67E-09
56*	3 (0)	23 (0)	27.4	72.9	113033041	1.06E-09	8.14E-09
57*	1 (0)	20 (0)	25.88	64.6	112275113	3.56E-10	7.13E-09
58*	3 (0)	15 (0)	33.13	90	113744025	1.06E-09	5.28E-09
59*	6 (0)	22 (0)	36.8	95.7	113232051	2.12E-09	7.77E-09
60	1 (1)	12 (8)	35.54	91.8	112982582	3.54E-10	4.25E-09
61	4 (4)	15 (11)	16.17	9.1	105115484	1.52E-09	5.71E-09
62	3 (0)	29 (1)	19.97	26.7	110479904	1.09E-09	1.05E-08
63*	5 (0)	13 (0)	16.95	12.8	107008993	1.87E-09	4.86E-09
64	7 (4)	15 (11)	28.85	77	113359158	2.47E-09	5.29E-09
65	3 (1)	16 (4)	24.33	53.9	112509267	1.07E-09	5.69E-09
66*	4 (0)	21 (0)	49.89	99.6	114602934	1.40E-09	7.33E-09
67	5 (0)	23 (1)	19.41	19.8	109643009	1.82E-09	8.39E-09
68*	2 (0)	25 (0)	54.38	99.7	114821644	6.97E-10	8.71E-09
69*	4 (0)	17 (0)	52.71	99.7	114761770	1.39E-09	5.93E-09
71	5(1)	13 (4)	24.43	53.8	112602641	1.78E-09	4.62E-09
72	2 (1)	12 (10)	41.13	97.5	114123223	7.01E-10	4.21E-09
73	4 (1)	22 (14)	45.75	98.8	114337770	1.40E-09	7.70E-09
74	7 (1)	17 (2)	44.65	99.1	114181513	2.45E-09	5.96E-09
75	5 (1)	23 (14)	37.96	94	113841913	1.76E-09	8.08E-09
76*	1 (0)	21 (0)	67.67	99.8	114923595	3.48E-10	7.31E-09
77	3 (1)	22 (4)	47.3	99	114412914	1.05E-09	7.69E-09
78*	2 (0)	19 (0)	44.09	98.4	113633870	7.04E-10	6.69E-09

MA Line	[a]Total indel	[a]Total SNM	Read depth coverage mean	Percentage of bases > 15x coverage	Accessible reference sites	Indel mutation rate	SNM mutation rate
79*	4 (0)	16 (0)	42.89	98.2	114042394	1.40E-09	5.61E-09
80*	4 (0)	20 (0)	55.83	99.7	114908544	1.39E-09	6.96E-09
81	3 (0)	20 (1)	64.52	99.8	115043934	1.04E-09	6.95E-09
82	6 (0)	18 (4)	49.21	99.5	114234723	2.10E-09	6.30E-09
83	0 (0)	21 (5)	67.38	99.8	115097888	0.00E+00	7.30E-09
84	1 (0)	16 (5)	50.2	99.1	114045897	3.51E-10	5.61E-09
85	3 (3)	16 (12)	39.96	95.4	113058882	1.06E-09	5.66E-09
86	1 (0)	22 (4)	28.75	71.5	112165524	3.57E-10	7.85E-09
88*	3 (0)	18 (0)	24.43	51.1	111214529	1.08E-09	6.47E-09
89*	2 (0)	20 (0)	22.15	41	110566748	7.24E-10	7.24E-09
91	9 (2)	16 (3)	36.85	93.9	113254326	3.18E-09	5.65E-09
92	6 (2)	27 (3)	41.04	94.5	113092246	2.12E-09	9.55E-09
94*	4 (0)	17 (0)	23.57	30.9	109824709	1.46E-09	6.19E-09
96	3 (0)	22 (4)	34.44	86.4	112885817	1.06E-09	7.80E-09
98*	4 (0)	21 (0)	24.61	43.9	110757184	1.44E-09	7.58E-09
99	1 (0)	25 (1)	31.74	76.6	111962083	3.57E-10	8.93E-09
100	3 (1)	12 (0)	11.38	1	73480966	1.63E-09	6.53E-09
101	1 (0)	23 (4)	13.85	3.3	94354608	4.24E-10	9.75E-09
102*	2 (0)	15 (0)	16.01	7.8	101580315	7.88E-10	5.91E-09
103	2 (1)	10 (0)	17.03	10.7	103397808	7.74E-10	3.87E-09
105*	4 (0)	15 (0)	54.17	99.7	115165246	1.39E-09	5.21E-09
106	3 (1)	21 (10)	33.66	90.1	114530293	1.05E-09	7.33E-09
108*	8 (0)	23 (0)	52.9	99.6	115126613	2.78E-09	7.99E-09
109*	4 (0)	20 (0)	62.49	99.8	115285394	1.39E-09	6.94E-09
110	5 (1)	24 (4)	39.92	94.7	114652928	1.74E-09	8.37E-09
111	2 (1)	18 (10	45.89	99.1	114977461	6.96E-10	6.26E-09
112*	6 (0)	24 (0)	48.33	99.3	115006139	2.09E-09	8.35E-09
113	3 (1)	18 (10)	37.29	92.4	114536331	1.05E-09	6.29E-09
115	3 (1)	21 (3)	31.7	82.2	114244406	1.05E-09	7.35E-09
114*	3 (0)	8 (0)	31.93	79.8	114185448	1.05E-09	2.80E-09
116	6 (1)	15 (3)	30.39	77	114106574	2.10E-09	5.26E-09
117	4 (0)	12 (2)	34.36	87.8	114395731	1.40E-09	4.20E-09
118	4 (0)	20 (2)	22.9	43.3	113004274	1.42E-09	7.08E-09
119*	6 (0)	26 (0)	27.28	64.5	113788727	2.11E-09	9.14E-09
Average mutation rate for the 48 MA lines (indicated by asterisks) that do not share mutations with other lines						1.30E-09	6.95E-09

[a]Number of shared mutation indicated in parenthesis

762 Supplementary Table S2. Recovered simulated reference mutations in five MA lines.

MA line	Accessible reference sites (proportion of whole genome)	Expected SNMs	Recovered SNMs
24	104239697 (87.5%)	87	86
31	90779408 (76.2%)	76	75
36	113252237 (95.1%)	95	93
62	112021390 (94%)	94	93
109	115724393 (97.1%)	97	98
Total		450	445

763

764 Supplementary Table S3. Recovered simulated sequence read mutations in five MA lines

MA line	Simulated homozygous SNMs	Filtered SNMs	Called heterozygous SNMs	Called homozygous SNMs
Read depth coverage > 0				
24	41	7	9	25 (61%)
31	61	12	9	40 (66%)
36	53	4	12	37 (70%)
62	61	7	7	47 (77%)
109	40	1	2	36 (90%)
Total	256	31	39	185 (72%)
Read depth coverage > 3				
24	39	7	7	25 (64%)
31	39	5	8	26 (67%)
36	23	2	3	18 (78%)
62	56	6	7	43 (77%)
109	39	1	2	36 (92%)
Total	196	21	27	148 (76%)
Read depth coverage > 10				
24	11	1	1	9 (82%)
31	0	0	0	Na
36	0	0	0	Na
62	19	1	2	16 (84%)
109	38	0	2	36 (95%)
Total	68	2	5	61 (90%)

765

766

767

768 Table S4. Comparison of the ratio of transitions to transversions at cytosines in MA lines.

	Transitions	Transversions	Transitions/Transversions	*p*-value[1]
Methylated cytosine	397	65	6.1	-
Non-methylated cytosine	604	120	5.0	0.252

769 [1]*p*-value of Fisher's Exact test

Table S5. Summary of logistic regression analysis for the effects of TEs, cytosine methylation, and chromosome regions on the likelihood of a given nucleotide being mutated in the MA lines. The base level is derived from non-TE, unmethylated bases and chromosome arm positions.

		Estimated natural log of odds	*p*-value	
Main effect	TE	1.80	2E-16	***
	Methylated	0.91	7E-10	***
	Pericentromeric/Centromere	0.57	2E-09	***
Two-way interaction	TE: Methylated	-0.14	0.574	
	TE: Pericentromeric/Centromere	-0.59	0.001	**
	Methylated: Pericentromeric/Centromere	0.30	0.116	
Three-way interaction	TE: Methylated: Pericentromeric/Centromere	-0.81	0.006	**

AIC = 94.2987. *p*-values < 0.01 (**) and <0.001 (***) are indicated.

775 Table S6. List of complex and double mutations in 107 MA lines.

MA line	Chromosome	Type	Position	Ref	Alt	Type
80	1	Double	13939374	T	A	SNM
			13939375	C	A	SNM
106	1	Double	15808720	G	A	SNM
			15808730	T	A	SNM
79	1	Double	16108568	A	T	SNM
			16108569	C	T	SNM
108	1	Double	19897266	T	G	SNM
			19897268	G	C	SNM
47	1	Double	29412132	G	A	SNM
			29412133	A	T	SNM
47	2	Double	1220067	G	A	SNM
			1220068	T	A	SNM
83	2	Double	2703908	T	A	SNM
			2703913	C	G	SNM
26	2	Double	6752832	G	T	SNM
			6752833	C	T	SNM
15	2	Double	12388974	T	C	SNM
			12388975	A	G	SNM
41	3	Double	3779277	G	A	SNM
			3779278	G	A	SNM
62	3	Double	6336607	G	A	SNM
			6336608	G	A	SNM
28	3	Double	12185747	T	A	SNM
			12185748	C	T	SNM

47	3	Double	12255447	G	A	SNM
			12255448	G	T	SNM
62	3	Double	14982296	C	T	SNM
			14982297	C	T	SNM
49	3	Double	17363410	C	T	SNM
			17363411	G	A	SNM
18	4	Double	3969772	T	A	SNM
			3969773	G	T	SNM
76	4	Double	4483683	A	T	SNM
			4483688	C	T	SNM
119	4	Double	13514556	G	T	SNM
			13514557	C	G	SNM
43	4	Double	17983901	A	C	SNM
			17983907	C	A	SNM
52	5	Double	898588	A	G	SNM
			898594	A	T	SNM
67	5	Double	1453729	T	A	SNM
			1453737	G	A	SNM
42	5	Double	3974068	G	A	SNM
			3974069	C	G	SNM
82	5	Double	13206438	G	A	SNM
			13206439	G	A	SNM
108	1	Complex	22447591	GTTACCAATCAGATCCATTA	G	Deletion
			22447614	GATT	G	Deletion
			22447618	A	G	SNM
2	1	Complex	22826184	TCGCCGTTAACTTTC	T	Deletion
			22826201	TATC	T	Deletion
28	2	Complex	16537492	ACCG	A	Deletion

			16537496	ATATCGGCCATGGTACATTAACACAAATAAAAT	A	Deletion
52	3	Complex	12507013	CAG	C	Deletion
			12507017	A	AG	Insertion
7	4	Complex	2049750	T	TGGTTGCAAAGTA	Insertion
			2049753	A	C	SNM
			2049755	T	G	SNM
			2049759	G	T	SNM
32	4	Complex	3107260	G	GACCT	Insertion
			3107263	GA	G	Deletion
			3107265	TATC	T	Deletion
78	4	Complex	6991719	T	C	SNM
			6991736	T	C	SNM
			6991767	T	C	SNM
			6991773	C	T	SNM
			6991790	T	C	SNM
			6991809	G	A	SNM
63	4	Complex	13918941	CGAGATCGAAAGAGGCA	C	Deletion
			13919005	G	T	SNM
77	4	Complex	18248959	G	A	SNM
			18248965	T	A	SNM
			18248975	G	A	SNM
33	5	Complex	8221423	C	T	SNM
			8221424	AGATTCAG	A	Deletion
35	5	Complex	11611778	C	CTTTTTTTTTTTT	Insertion
			11611796	TGA	T	Deletion
36	5	Complex	11662553	C	CTTCAGT	Insertion
			11662558	A	T	SNM
82	5	Complex	13206428	T	A	SNM

			13206429	TC	T	Deletion
			13206434	T	G	SNM
			13206436	G	A	SNM
			13206438	G	A	SNM
119	5	Complex	13456657	T	C	SNM
			13456659	G	T	SNM
			13456667	T	G	SNM
			13456673	G	A	SNM
			13456794	T	A	SNM
75	5	Complex	18240081	TGGCAG	T	Deletion
			18240087	TGC	T	Deletion
			18240090	TAA	T	Deletion
3	5	Complex	20295146	T	TAG	Insertion
			20295149	TTC	T	Deletion

776