Line filtering:

Removed the following lines based on missing data:

Line %missing

128 27.40

ste 84.44

Sz116 75.95

Sz121 98.69

Sz158 97.69

Sz21 98.25

Sz286 97.84

md106ts 94.02

Lines with high heterozygosity:

Line %het

Sz99 9.18

Sz91 9.55

Sz49 12.23

Sz5 10.55

Sz293 12.59

Sz283 10.01

Sz265 16.40

Sz261 12.77

Sz246 9.78

Sz244 10.85

Sz243 9.14

Sz237\* 34.16\*

Sz218 19.02

Sz208 9.07

Sz209 9.01

Sz19 14.75

Sz11 11.86

\*Removed from analysis

Total loci: **7,305,030**

Bialleles only: **6,921,236**

Bialleles only, not contigs: **6,755,185**

Loci on contigs: **166,051**

Non-biallelic loci: **383,794**

Sites with NO missing data: **1,513,147**

Loci filtering

If more than n% lines have missing data at a position:

Percent missing Sites retained Percent of total

90% (>167 lines) **6,616,640 97.95%**

75% (>140 lines) **6,594,921 97.62%**

50% (>93 lines) **6,527,945 96.63%**

25% (>46 lines) **6,378671 94.42%**

10% (>20 lines) **5,998,575 88.79%**