

the investigation of 149-snp peak data

Yu Huang

November 20, 2007

Abstract

look at the peak data for different calls

Contents

1 cluster plots	1
1.1 cluster plots for single snp	12

1 cluster plots

282802 out of 955388 snp calls (30%) has raw peak(intensity) data. check their cluster plots.

For every snp genotyped, there're two alleles. Sequenom uses mass spectrometer to distinguish different alleles. It produces the intensity of peak1 and peak2, which corresponds to allele 1 and 2 respectively.

A cluster plot (name comes from the supplementary of the Wellcome Trust 2007 paper. <https://dl324b-1.cmb.usc.edu/research/ref/nobody2007>) is a graphical representation of the results of both the genotyping and genotype calling of a SNP. It is a scatter plot of the two peak intensities from the genotyping, with each point representing one snp calling in one ecotype/strain. Each point is colored according to how genotype calling algorithm decided to classify that particular snp in one ecotype. 4 different classifications are homozygote for allele 1 (red), homozygote for allele 2 (blue), heterozygote for two alleles (green), missing/NA (yellow).

Figure 1-7 are cluster plots with all snps concatenated together.

Figure 11 and onwards are the cluster plots for one single snp (ordered by chromosome, position). the genotype calling algorithm might have different parameters for different snps.

One apparent problem is that same color dots appear on two off-diagonal sides. This can be seen in some single-snp cluster plots. (2007-10-30) Figure 2 and 1 shows this issue is not as serious for allele 2 homozygote as for allele 1 homozygote. it looks like it doesn't happen at all to allele 2 homozygote calling (dots all above the diagonal line in Figure 2). this raises the suspicion that the sequenom algorithm uses the diagonal line to separate data. this diagonal cut is also visible in Figure 1.

Three Possibilities:

1. some peak1/peak2 data was swapped in the database.
2. the sequenom genotype calling algorithm would be very lousy by calling some points from both sides of the heterozygous cloud as the same allele.

3. the correponding allele for peak1/peak2 is changing even for the same snp.

we need to check this to get a solid judgement.

but for some snps, like Figure 19, 35, 40, 53, 63, 65, 66, 78, 99, 105, 107, 119, 123, 130, etc, the cluster plots don't look good. some of them might be good after we figure out the cause of same-color-both-side problem. they might correspond to snps with too many hets , or too many NAs.

(2007-10-30) one more thing is dots sitting on the X or Y axis. One of the peak1/peak2 is 0. is that real 0? given the noise in MS(mass spectrometry), it's hard to get absolute 0, isn't it? i'm no MS expert, willing to hear feedback.

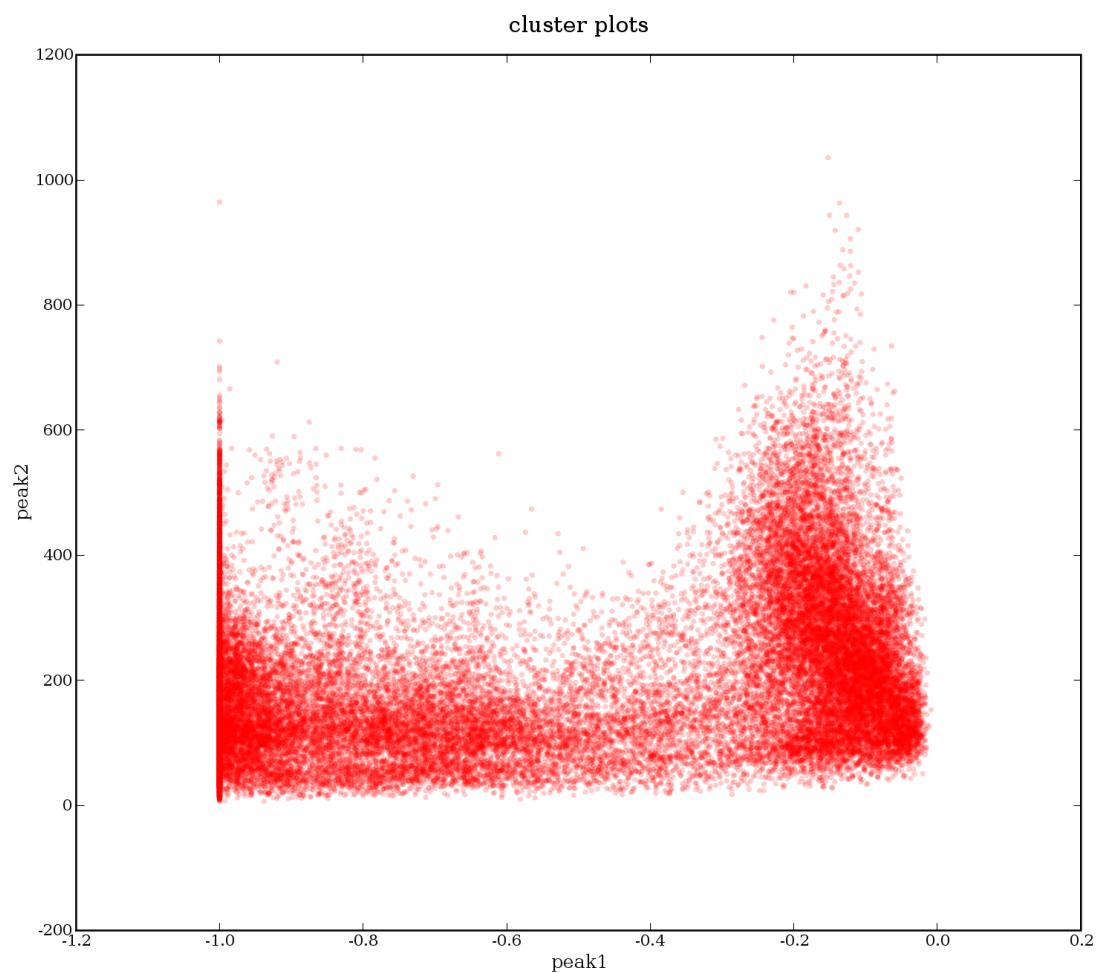


Figure 1: homozygote for allele 1

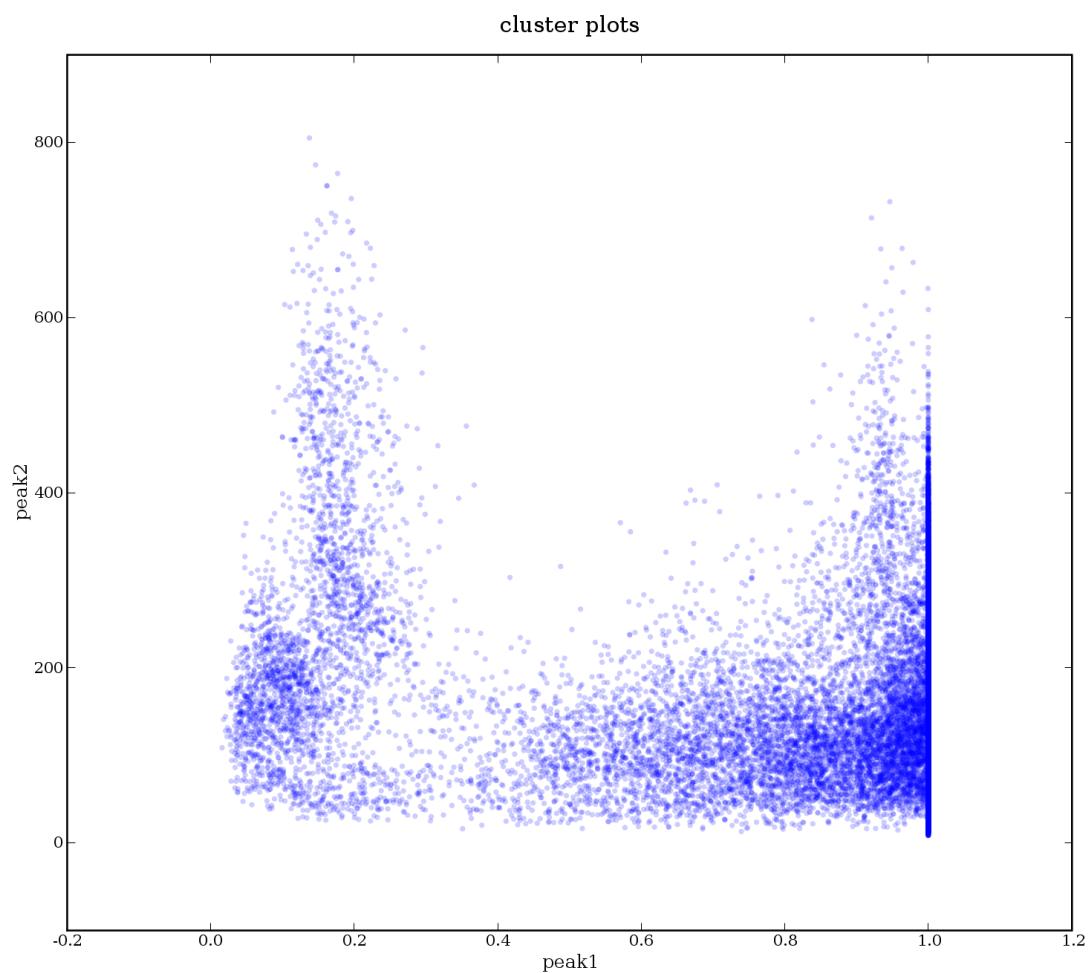


Figure 2: homozygote for allele 2

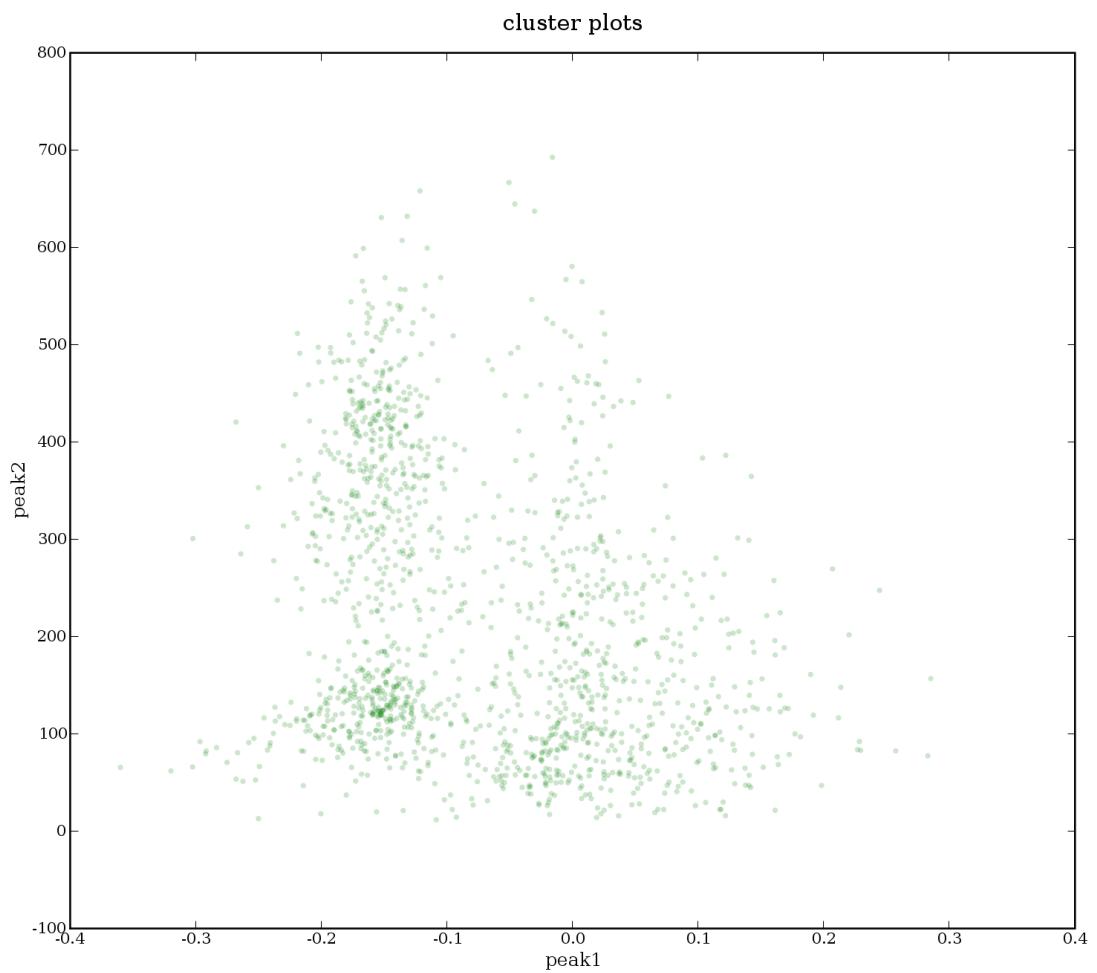


Figure 3: heterozygote

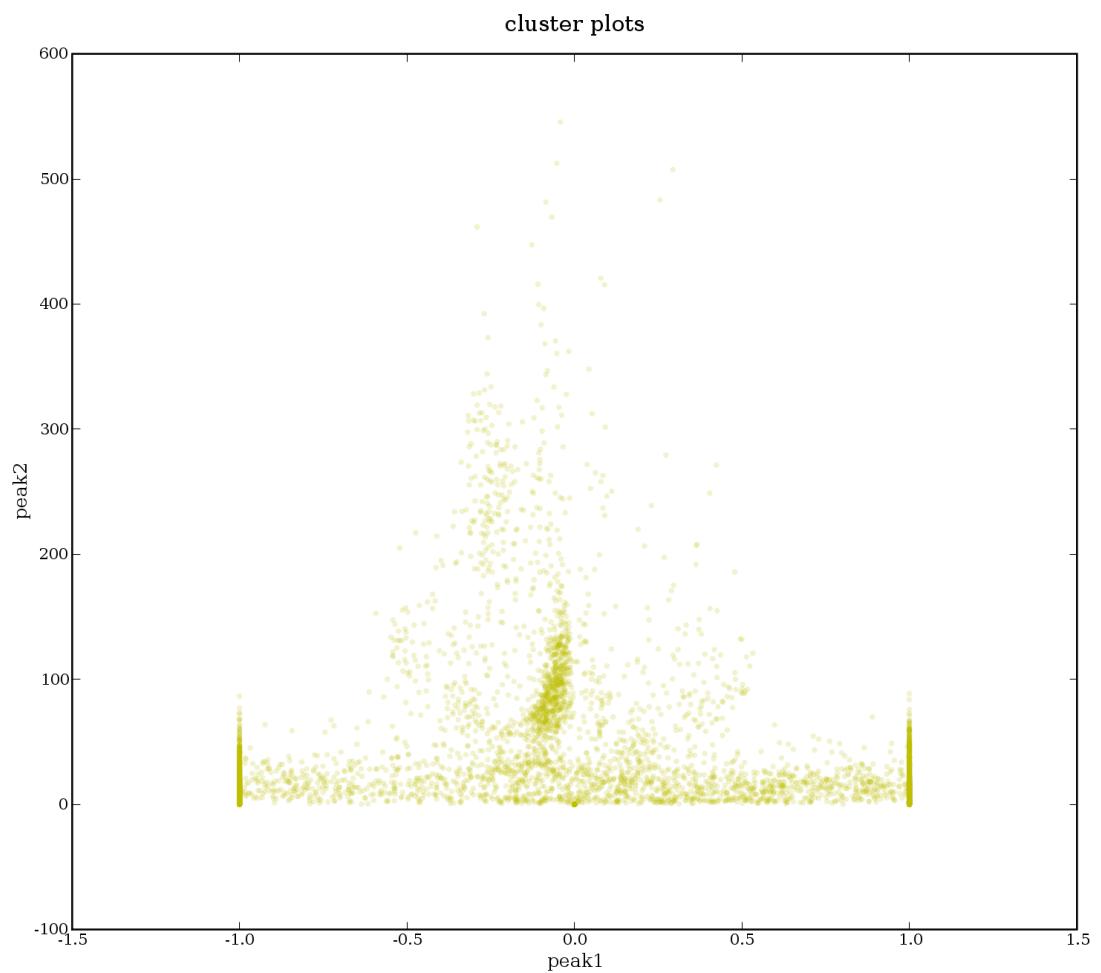


Figure 4: missing/NA

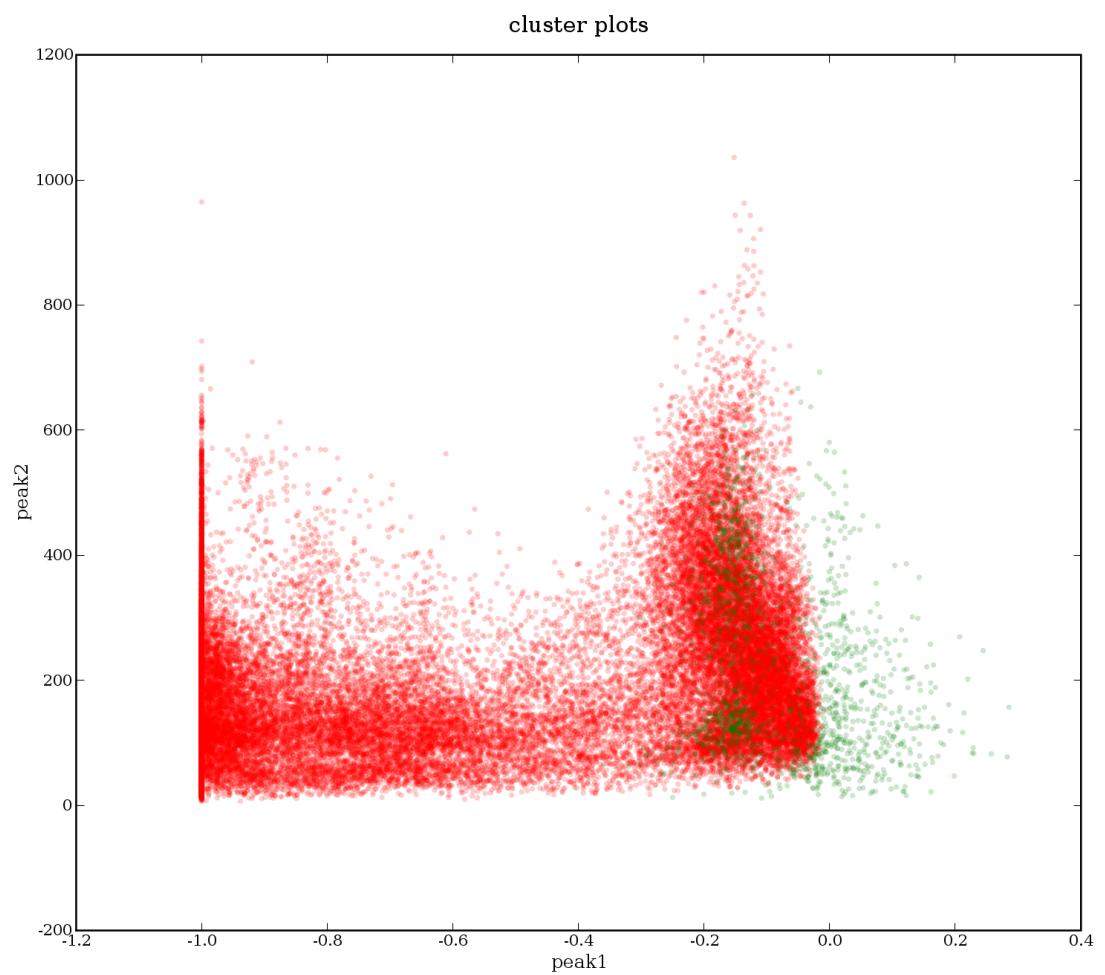


Figure 5: overlapping of homozygote for allele 1 and heterozygote

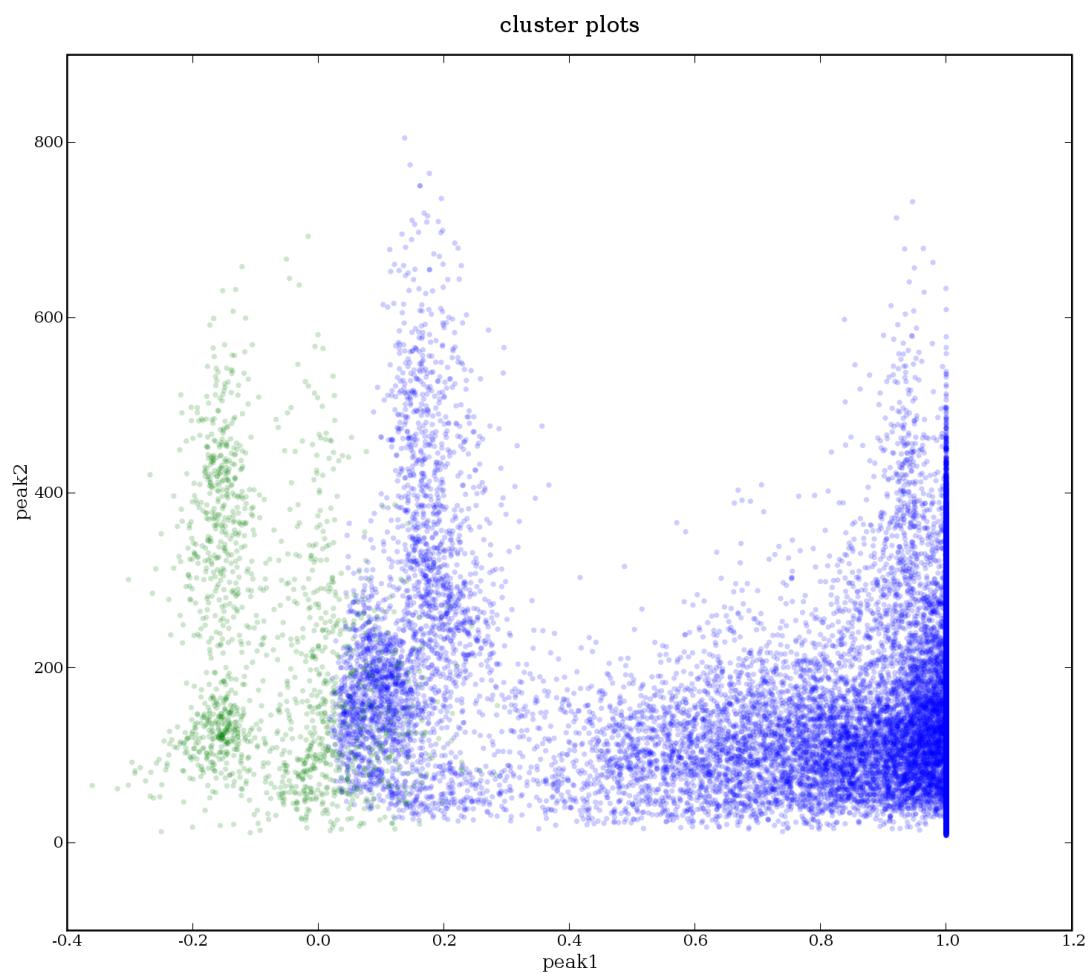


Figure 6: overlapping of homozygote for allele 2 and heterozygote

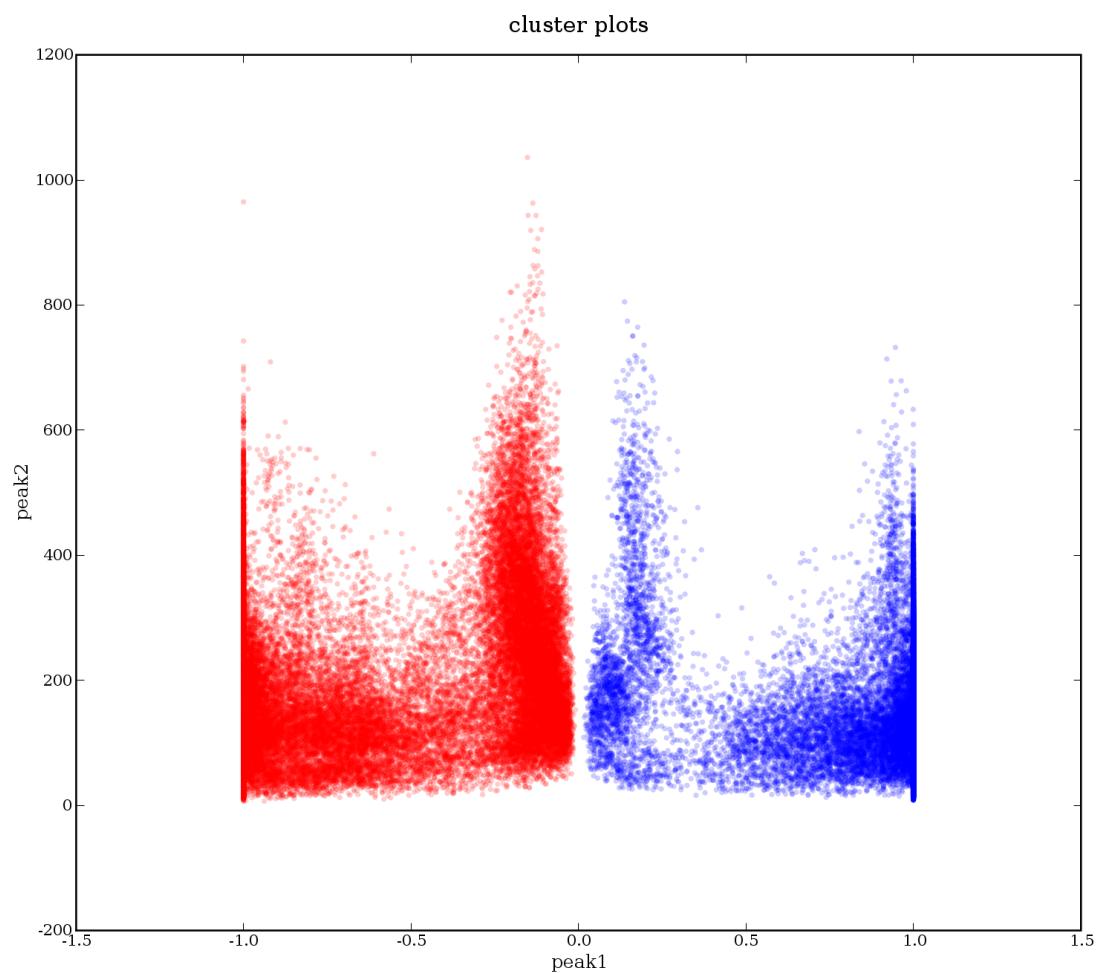


Figure 7: overlapping of homozygote for allele 1 and allele 2

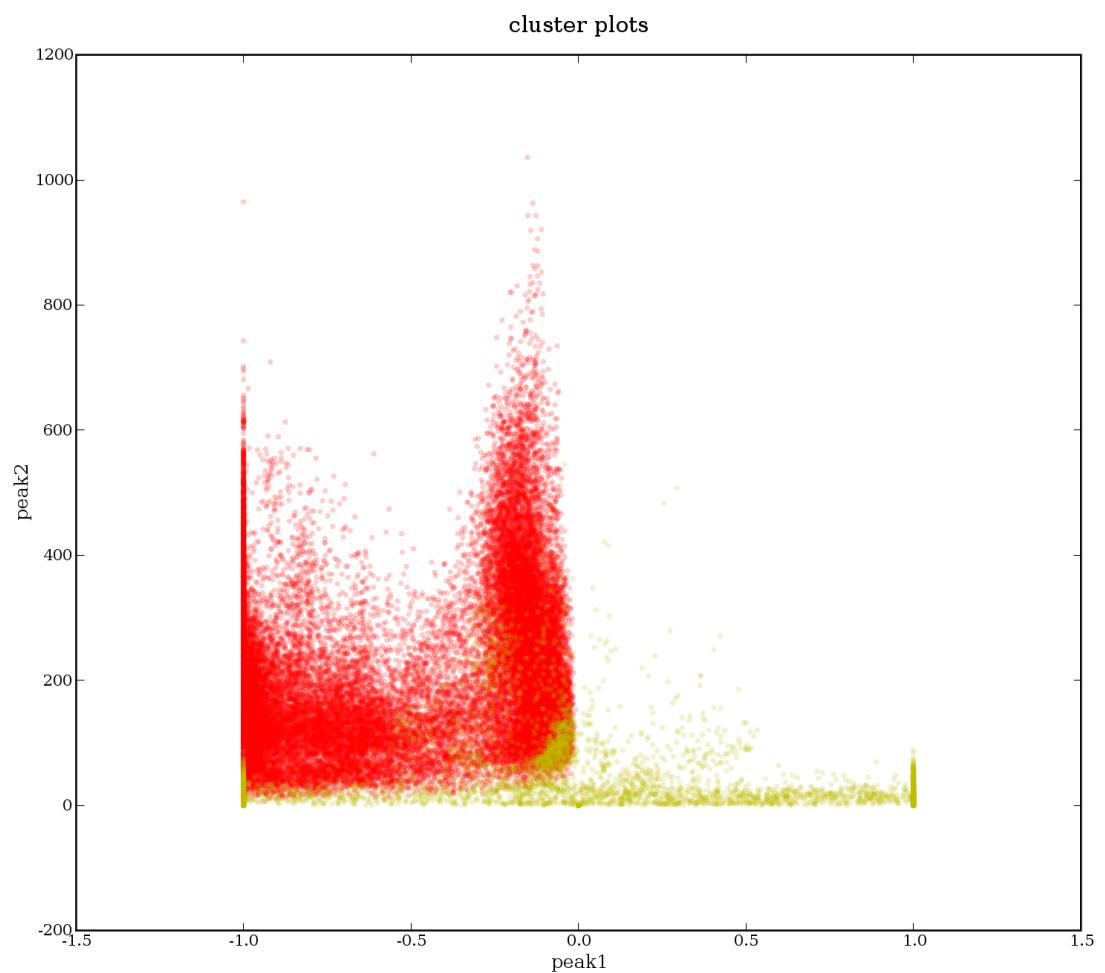


Figure 8: overlapping of homozygote for allele 1 and NA

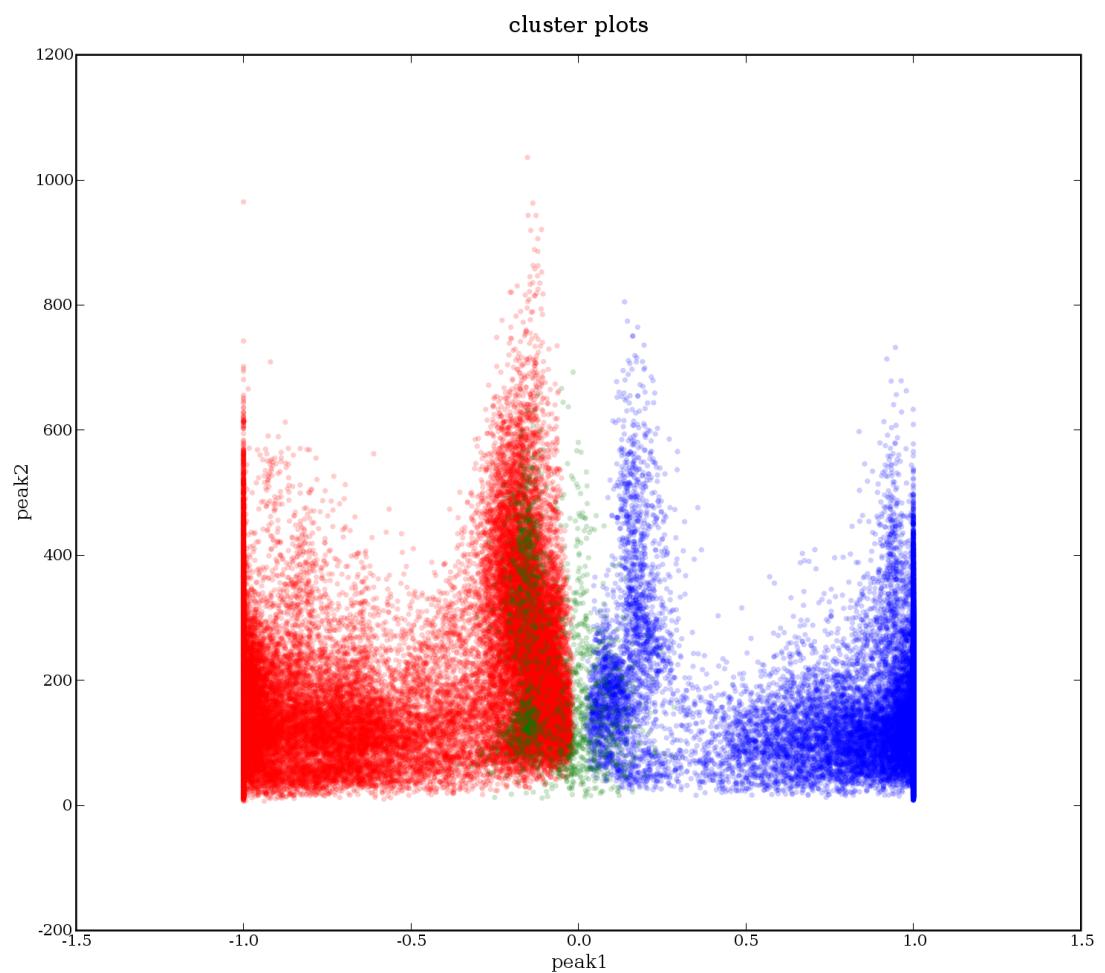


Figure 9: homozygote and heterozygote

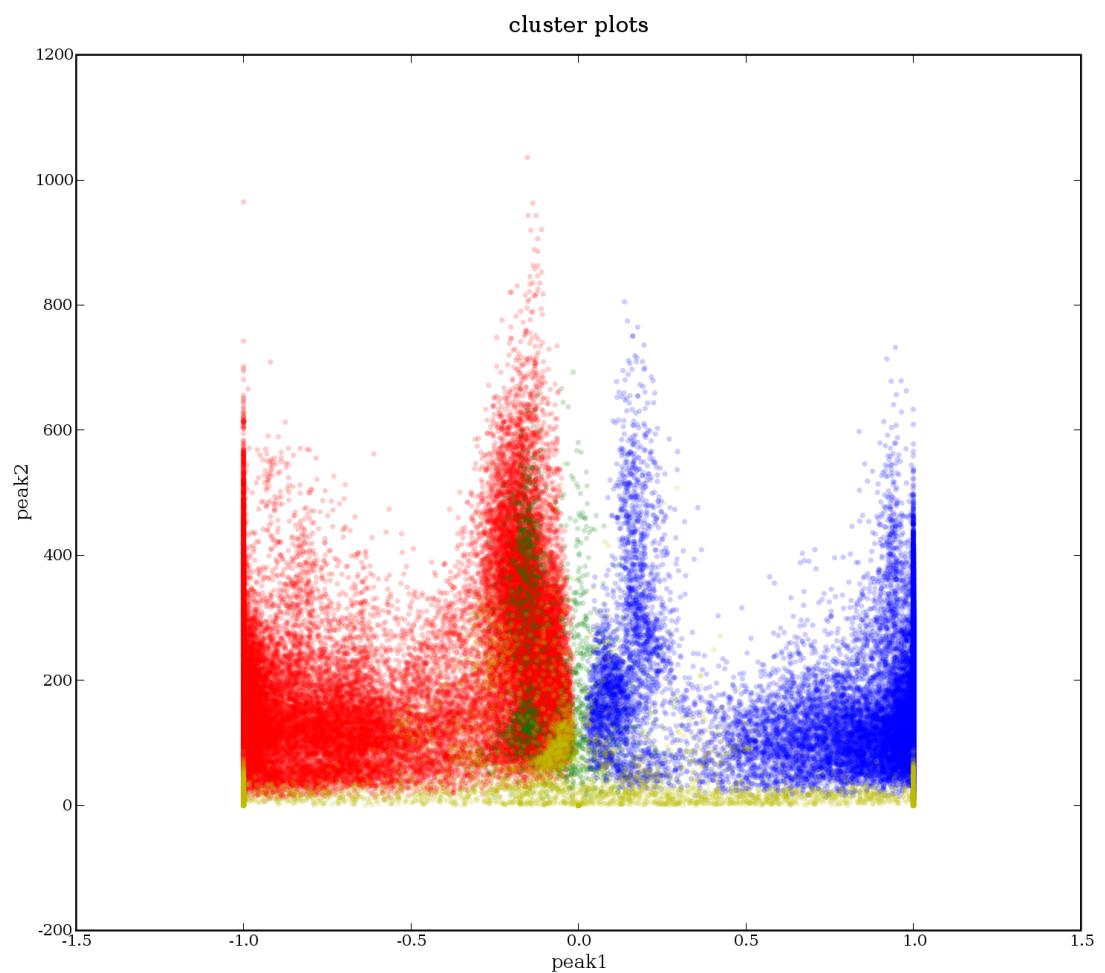


Figure 10: all four different calls

1.1 cluster plots for single snp

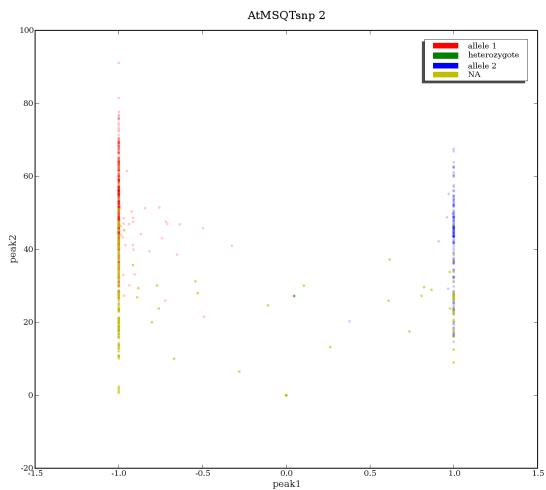


Figure 11: cluster plot for AtMSQTsnp 2.

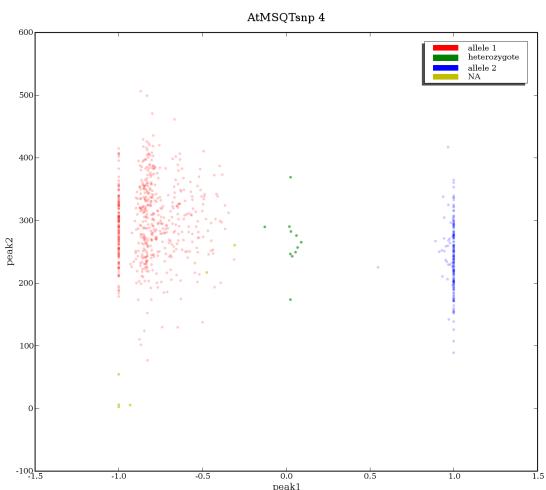


Figure 12: cluster plot for AtMSQTsnp 4.

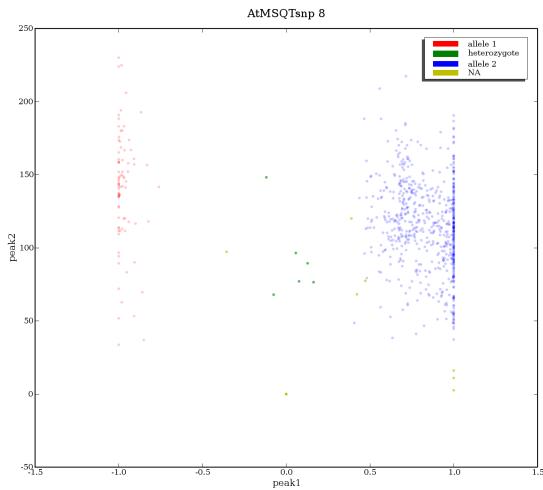


Figure 13: cluster plot for AtMSQTsnp 8.

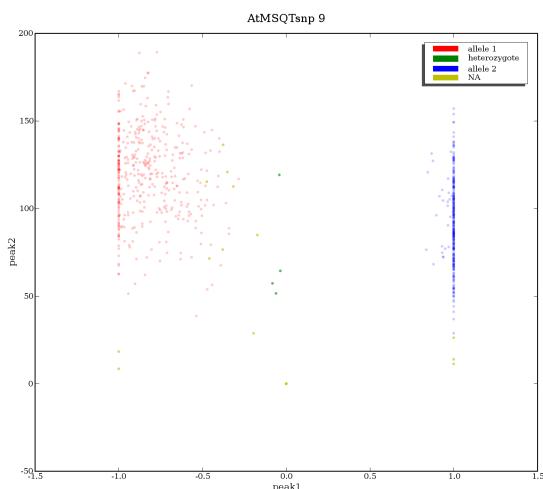


Figure 14: cluster plot for AtMSQTsnp 9.

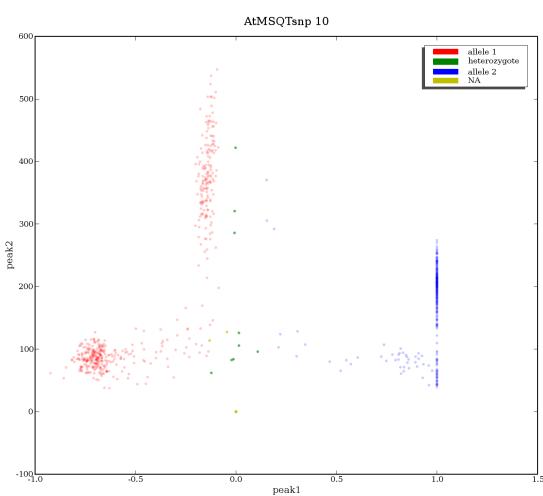


Figure 15: cluster plot for AtMSQTsnp 10.

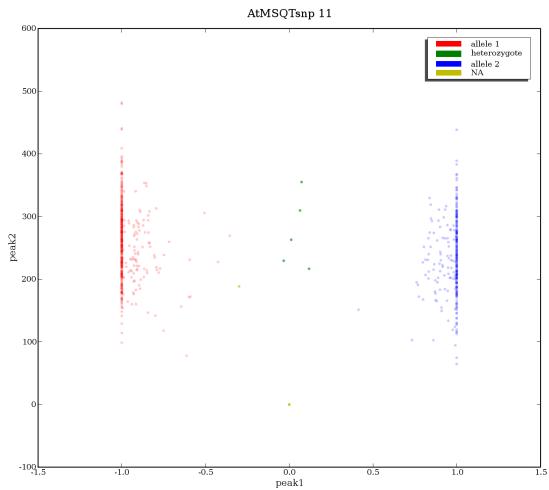


Figure 16: cluster plot for AtMSQTsnp 11.

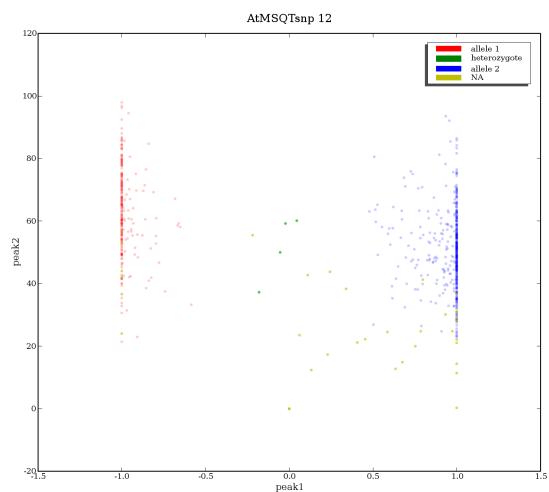


Figure 17: cluster plot for AtMSQTsnp 12.

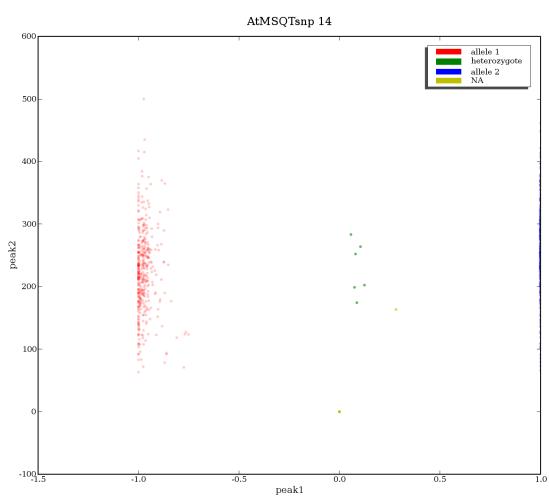


Figure 18: cluster plot for AtMSQTsnp 14.

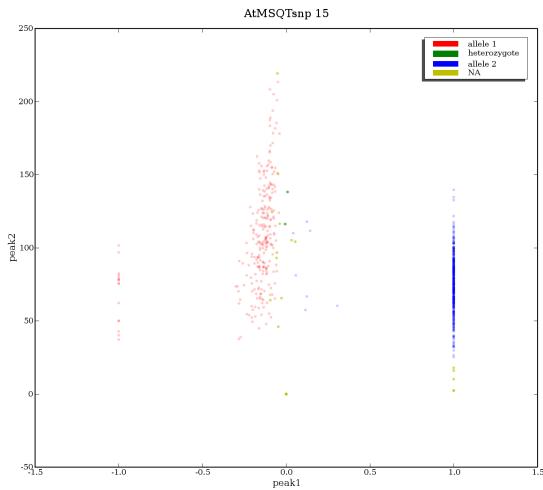


Figure 19: cluster plot for AtMSQTsnp 15.

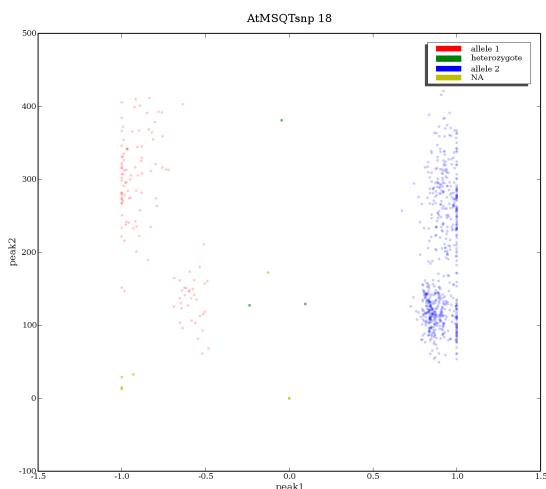


Figure 20: cluster plot for AtMSQTsnp 18.

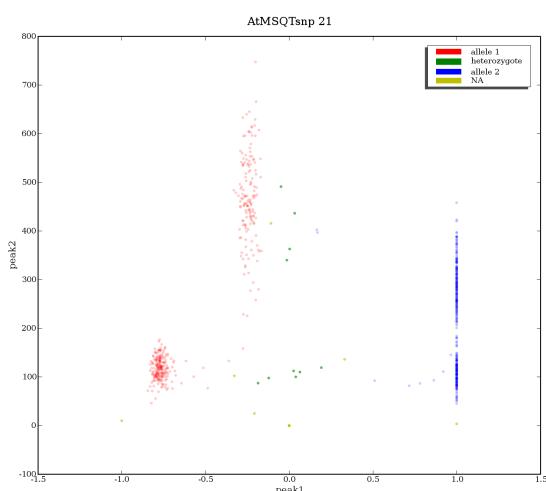


Figure 21: cluster plot for AtMSQTsnp 21.

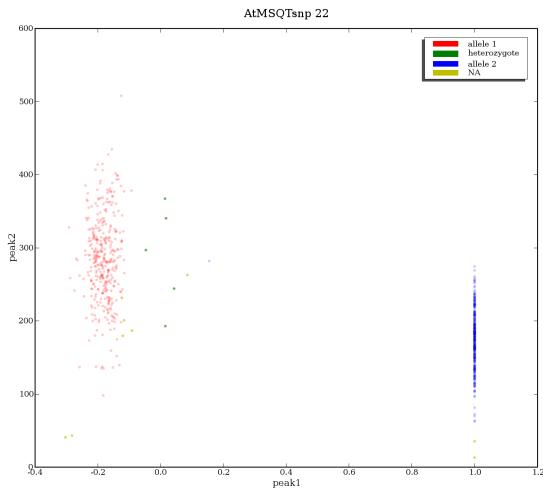


Figure 22: cluster plot for AtMSQTsnp 22.

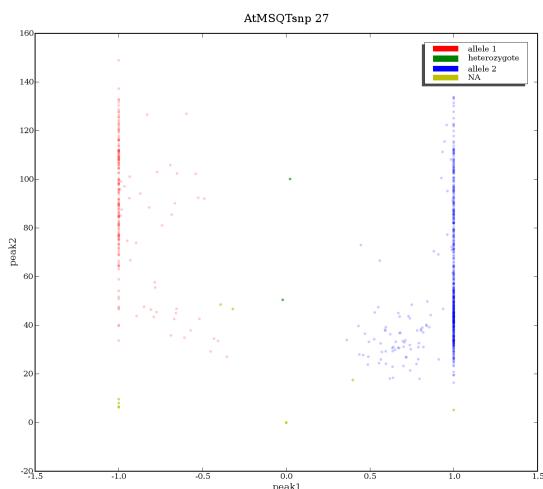


Figure 23: cluster plot for AtMSQTsnp 27.

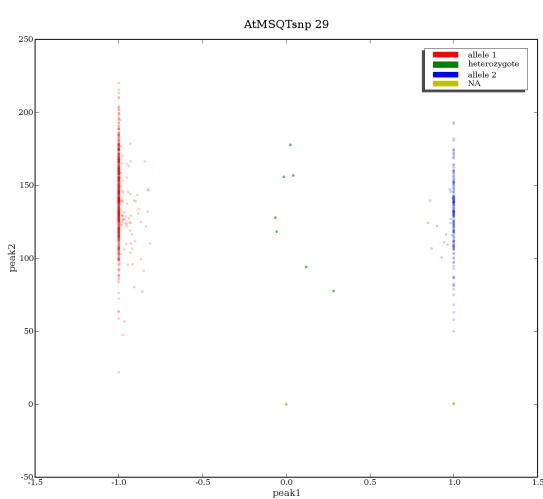


Figure 24: cluster plot for AtMSQTsnp 29.

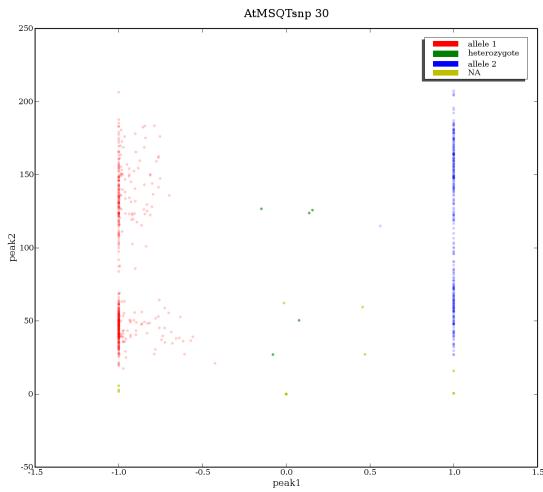


Figure 25: cluster plot for AtMSQTsnp 30.

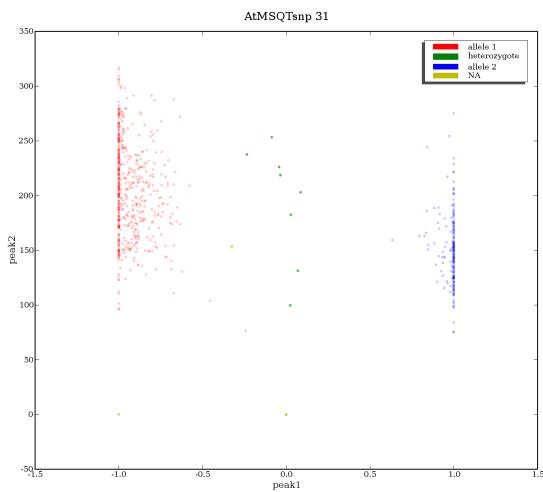


Figure 26: cluster plot for AtMSQTsnp 31.

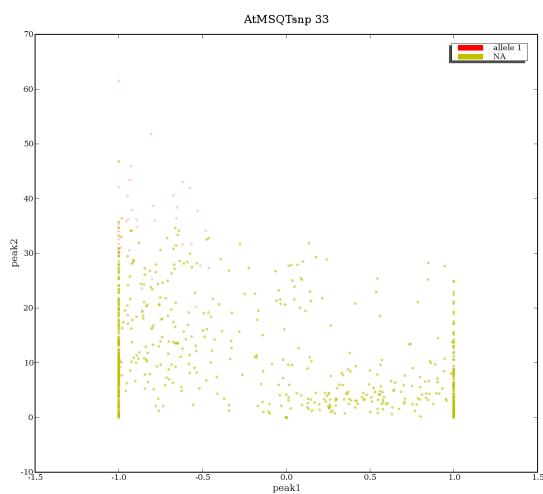


Figure 27: cluster plot for AtMSQTsnp 33.

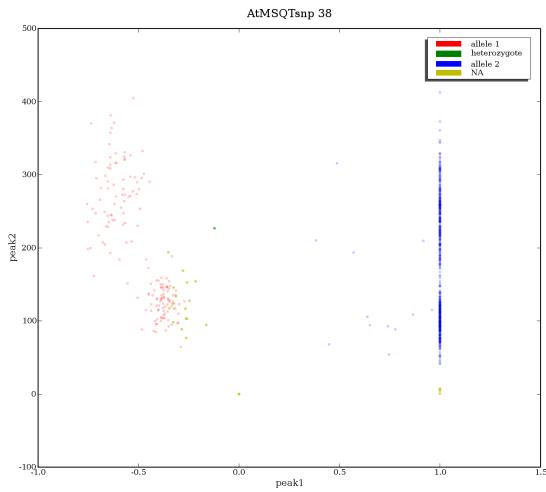


Figure 28: cluster plot for AtMSQTsnp 38.

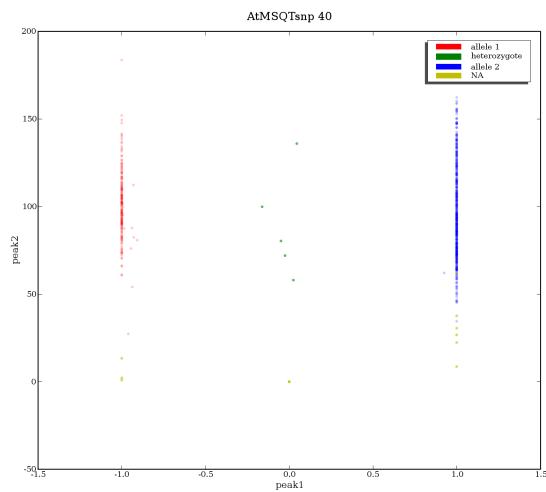


Figure 29: cluster plot for AtMSQTsnp 40.

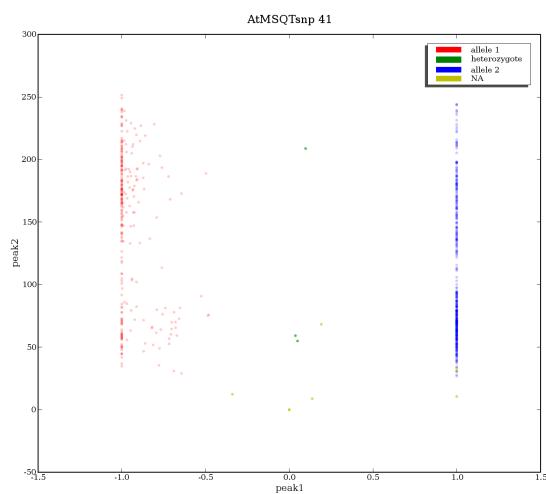


Figure 30: cluster plot for AtMSQTsnp 41.

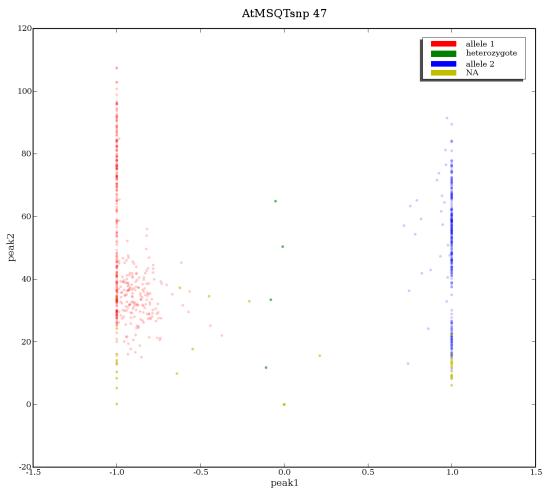


Figure 31: cluster plot for AtMSQTsnp 47.

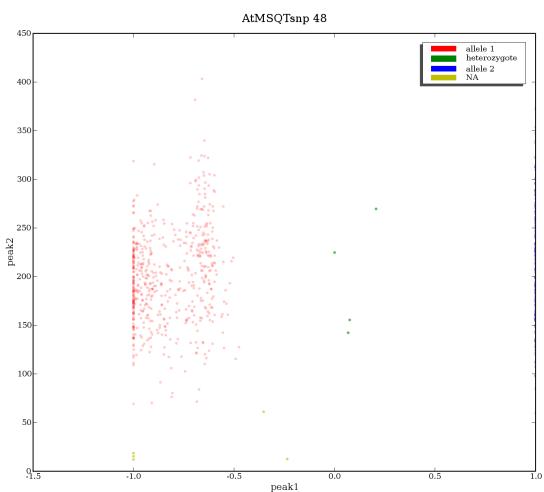


Figure 32: cluster plot for AtMSQTsnp 48.

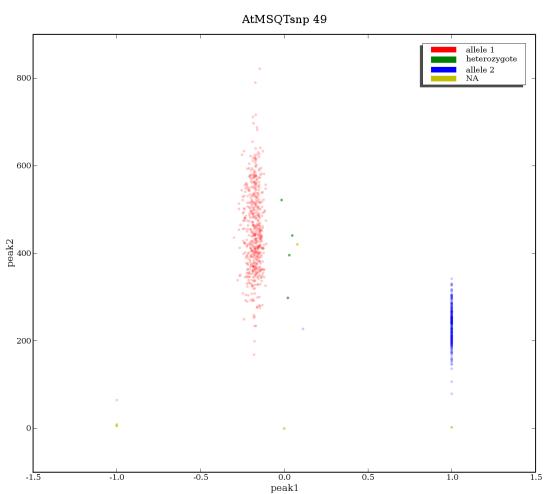


Figure 33: cluster plot for AtMSQTsnp 49.

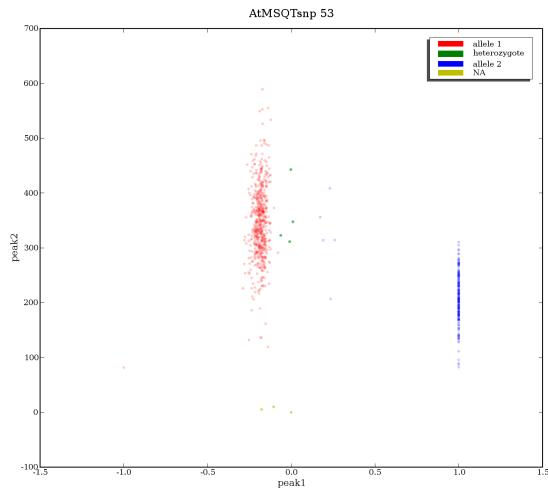


Figure 34: cluster plot for AtMSQTsnp 53.

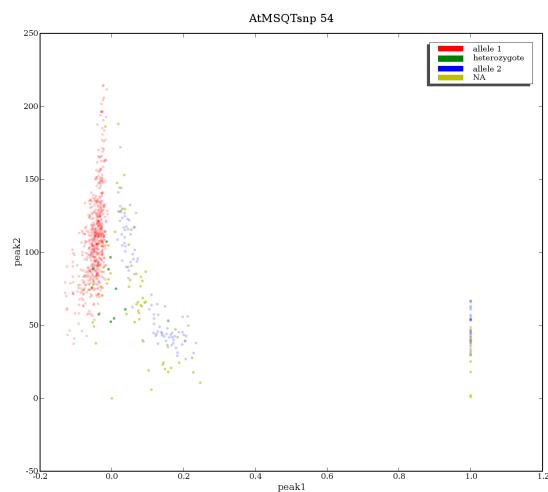


Figure 35: cluster plot for AtMSQTsnp 54.

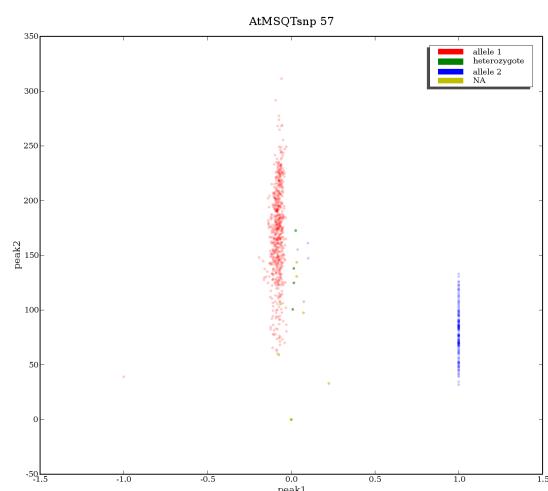


Figure 36: cluster plot for AtMSQTsnp 57.

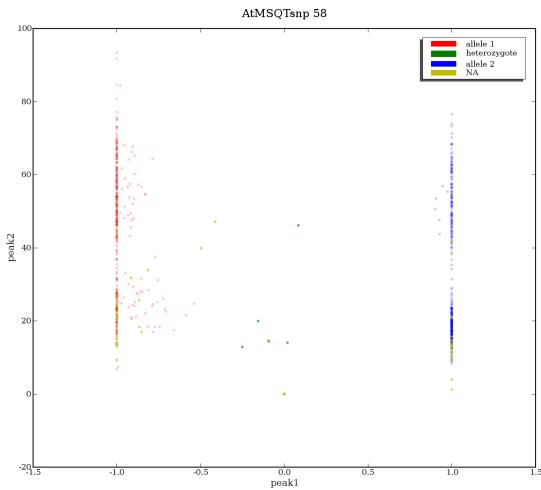


Figure 37: cluster plot for AtMSQTsnp 58.

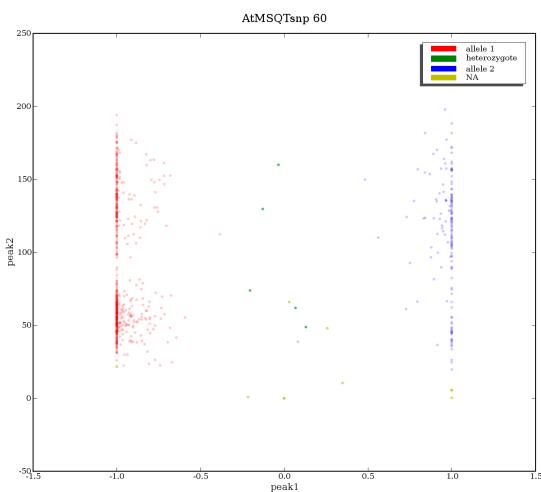


Figure 38: cluster plot for AtMSQTsnp 60.

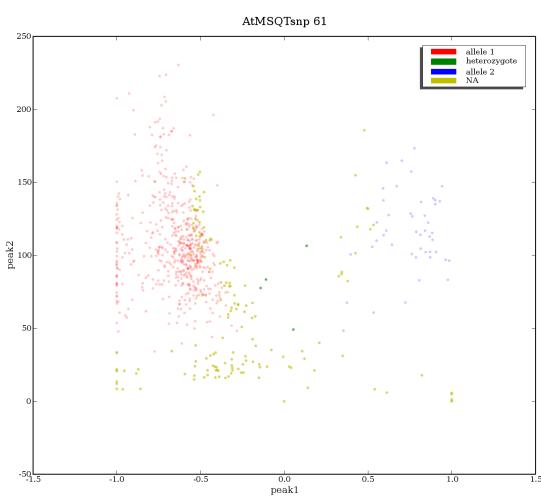


Figure 39: cluster plot for AtMSQTsnp 61.

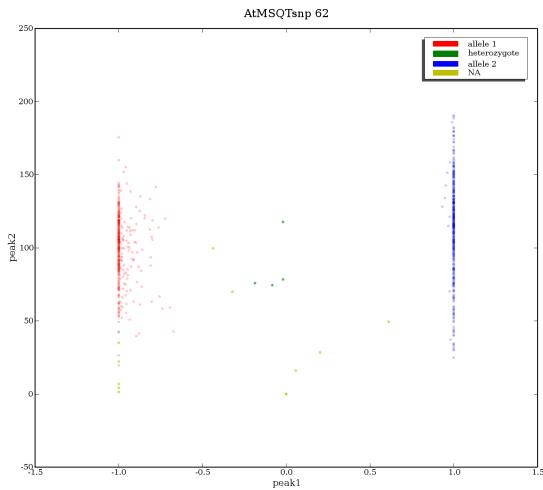


Figure 40: cluster plot for AtMSQTsnp 62.

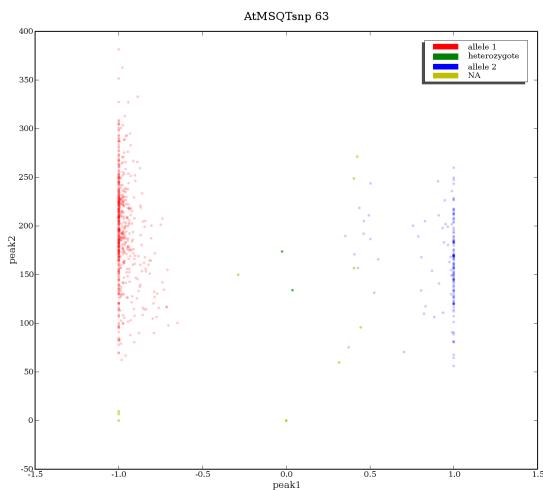


Figure 41: cluster plot for AtMSQTsnp 63.

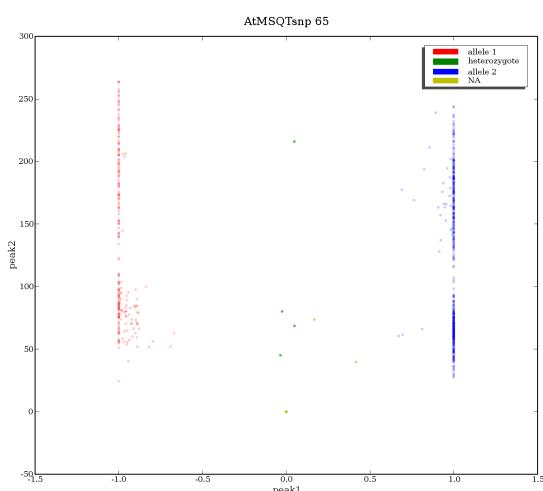


Figure 42: cluster plot for AtMSQTsnp 65.

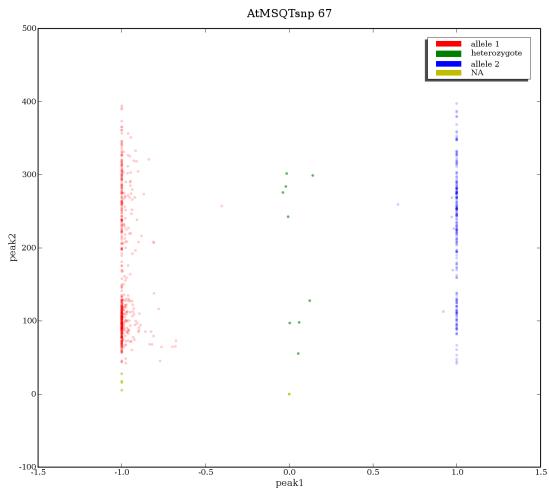


Figure 43: cluster plot for AtMSQTsnp 67.

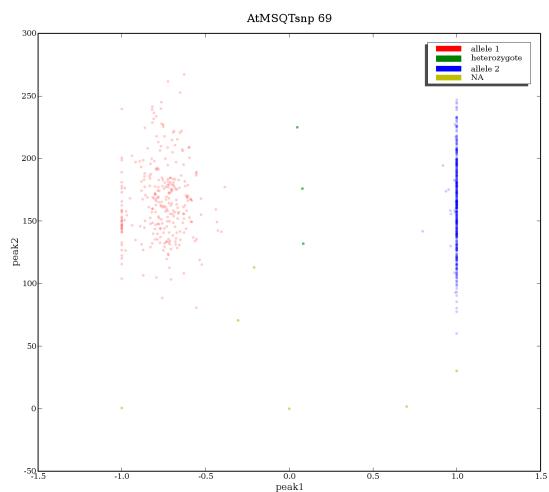


Figure 44: cluster plot for AtMSQTsnp 69.

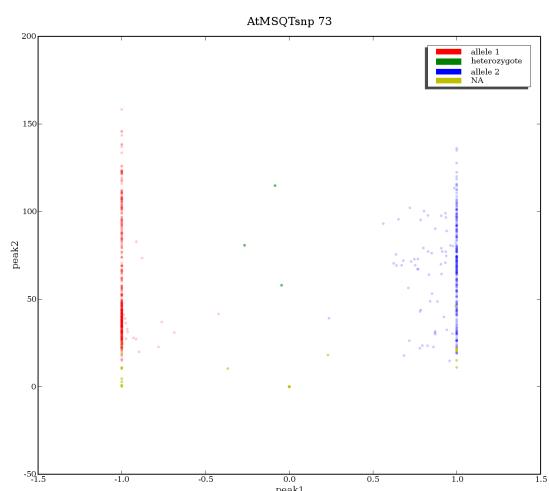


Figure 45: cluster plot for AtMSQTsnp 73.

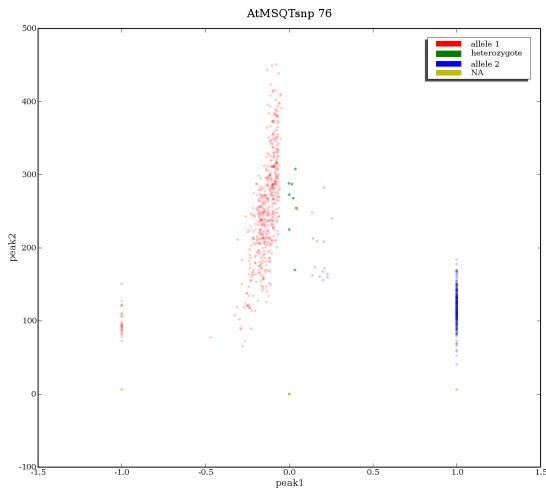


Figure 46: cluster plot for AtMSQTsnp 76.

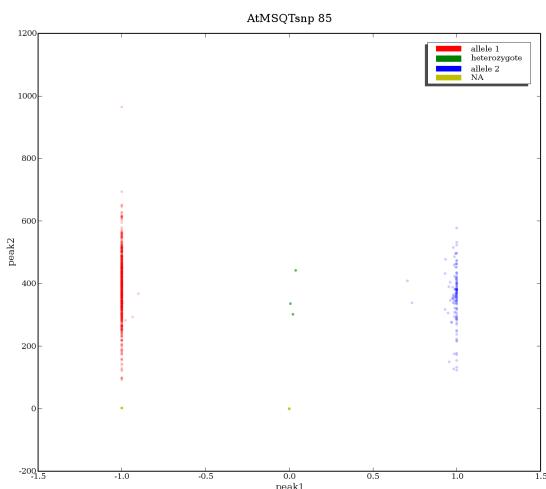


Figure 47: cluster plot for AtMSQTsnp 85.

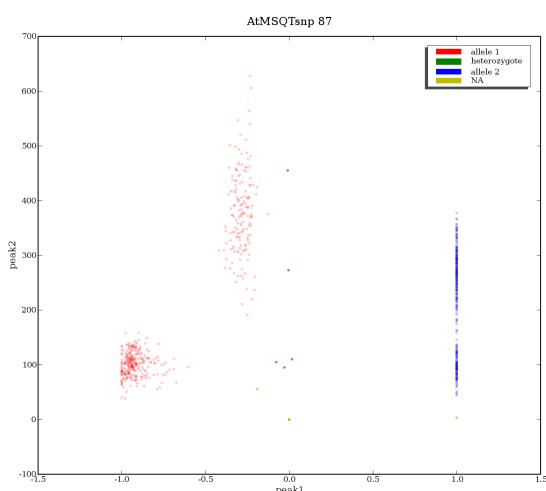


Figure 48: cluster plot for AtMSQTsnp 87.

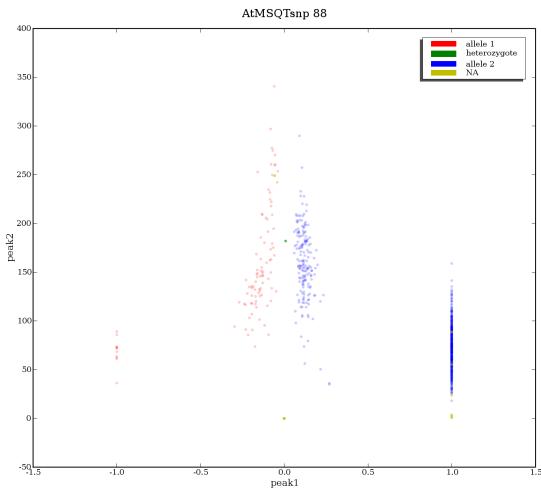


Figure 49: cluster plot for AtMSQTsnp 88.

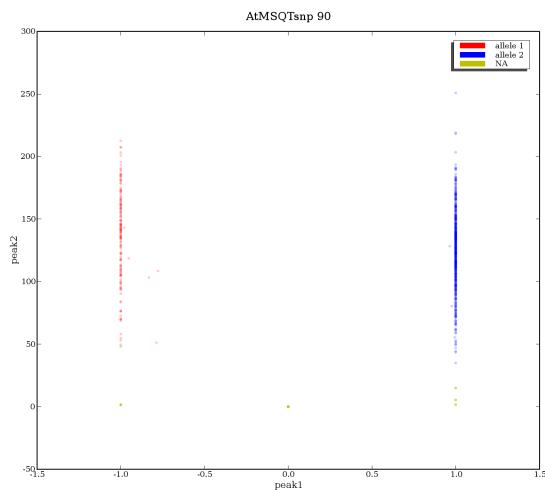


Figure 50: cluster plot for AtMSQTsnp 90.

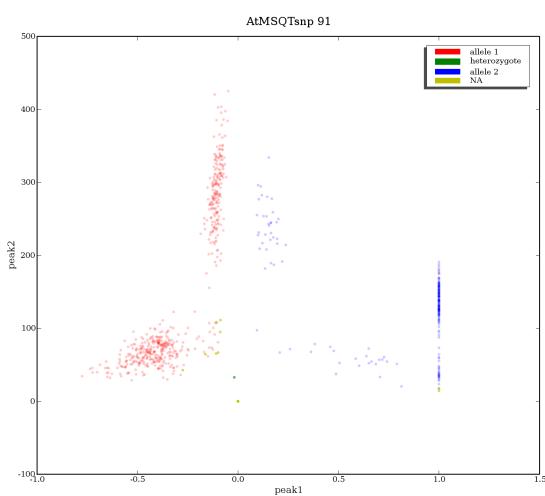


Figure 51: cluster plot for AtMSQTsnp 91.

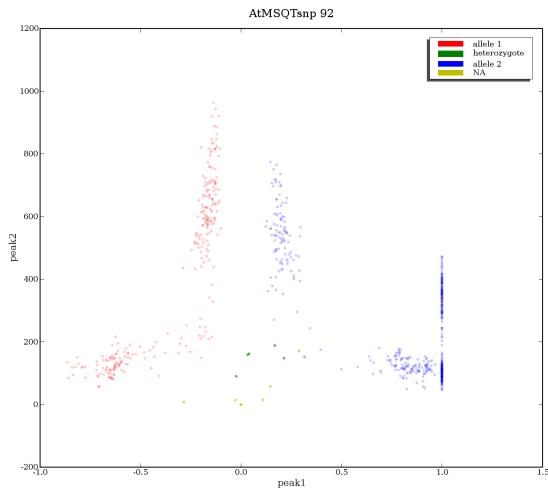


Figure 52: cluster plot for AtMSQTsnp 92.

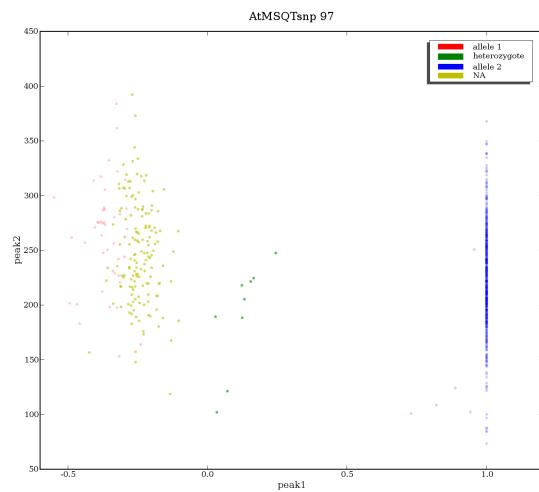


Figure 53: cluster plot for AtMSQTsnp 97.

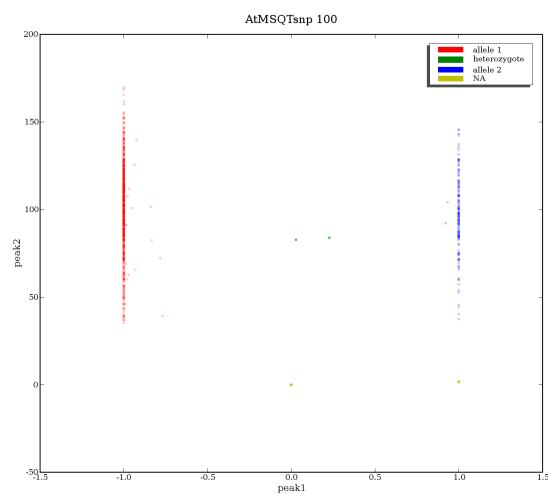


Figure 54: cluster plot for AtMSQTsnp 100.

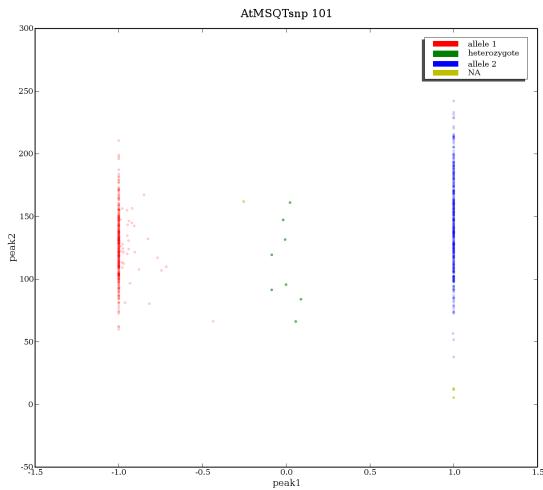


Figure 55: cluster plot for AtMSQTsnp 101.

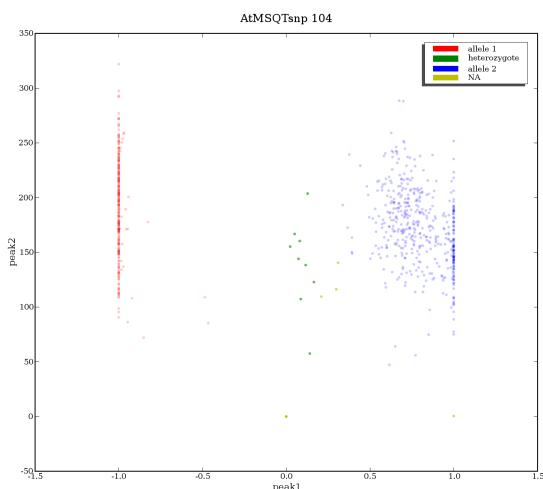


Figure 56: cluster plot for AtMSQTsnp 104.

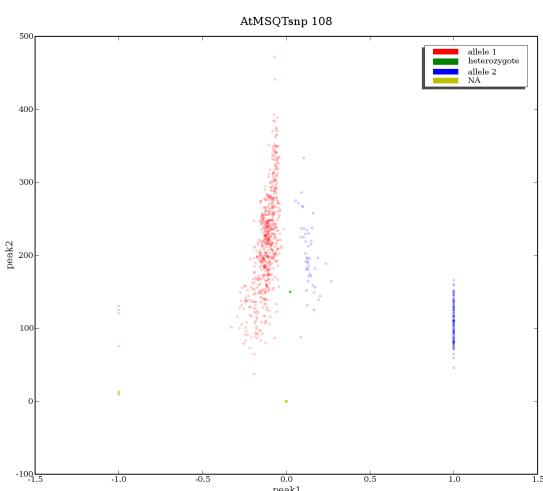


Figure 57: cluster plot for AtMSQTsnp 108.

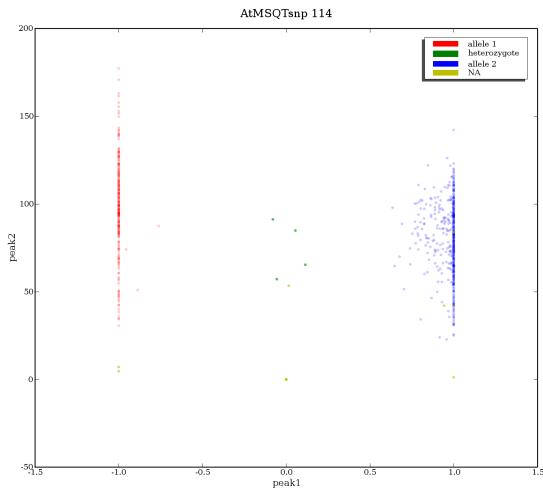


Figure 58: cluster plot for AtMSQTsnp 114.

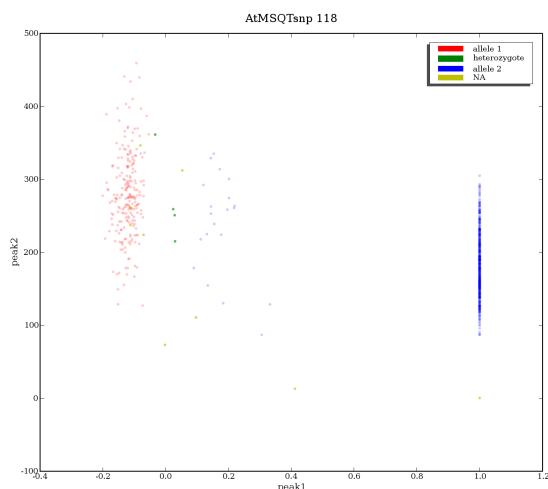


Figure 59: cluster plot for AtMSQTsnp 118.

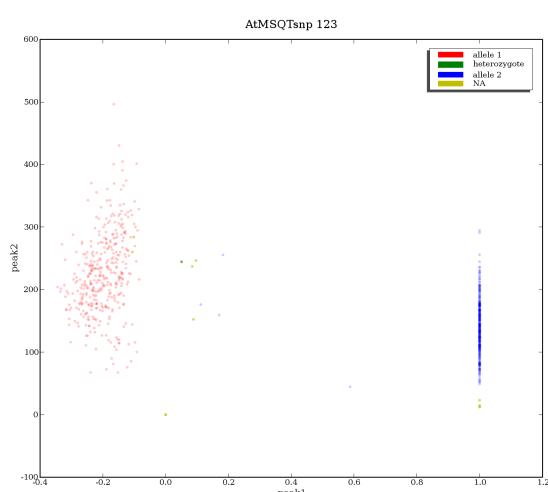


Figure 60: cluster plot for AtMSQTsnp 123.

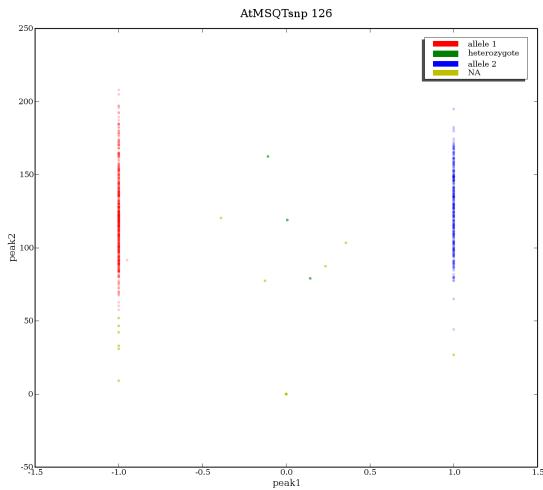


Figure 61: cluster plot for AtMSQTsnp 126.

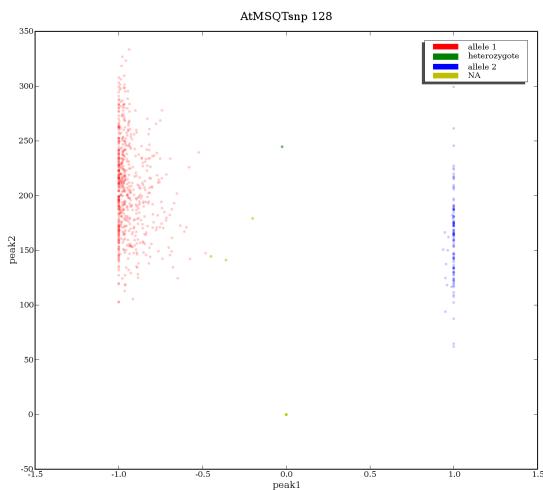


Figure 62: cluster plot for AtMSQTsnp 128.

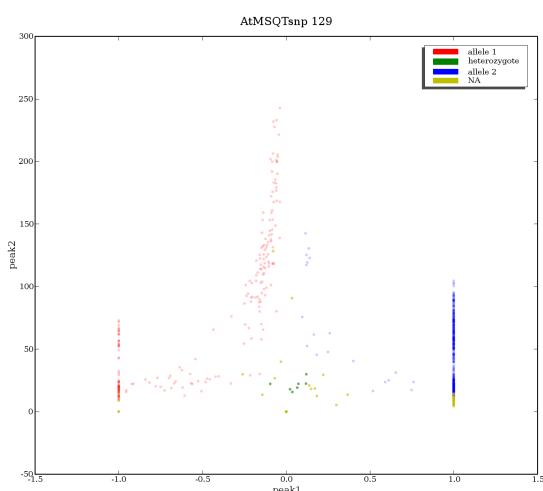


Figure 63: cluster plot for AtMSQTsnp 129.

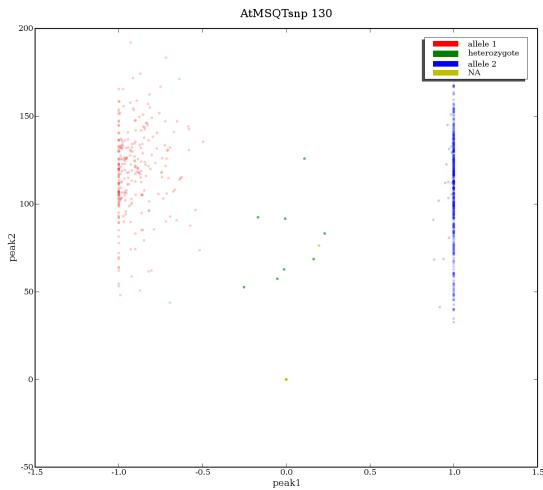


Figure 64: cluster plot for AtMSQTsnp 130.

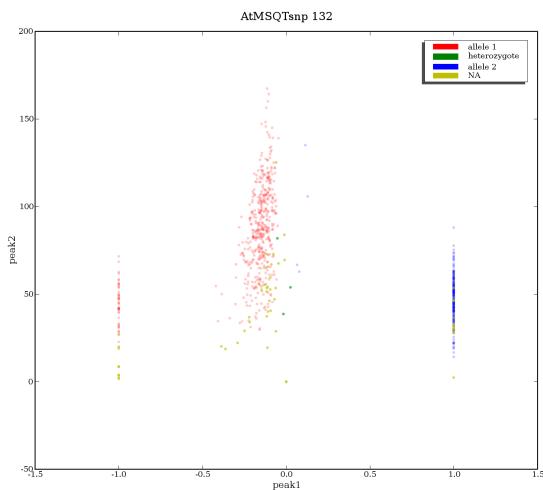


Figure 65: cluster plot for AtMSQTsnp 132.

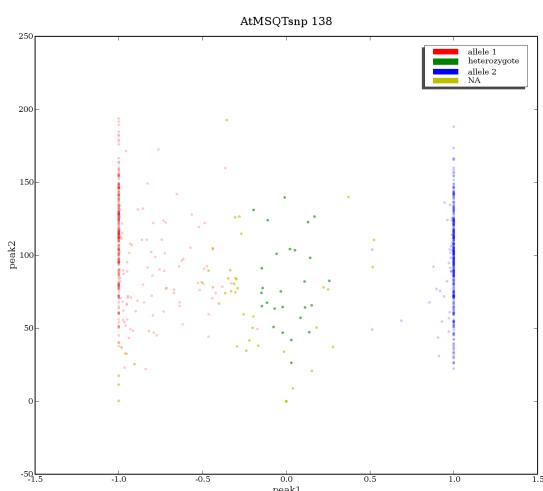


Figure 66: cluster plot for AtMSQTsnp 138.

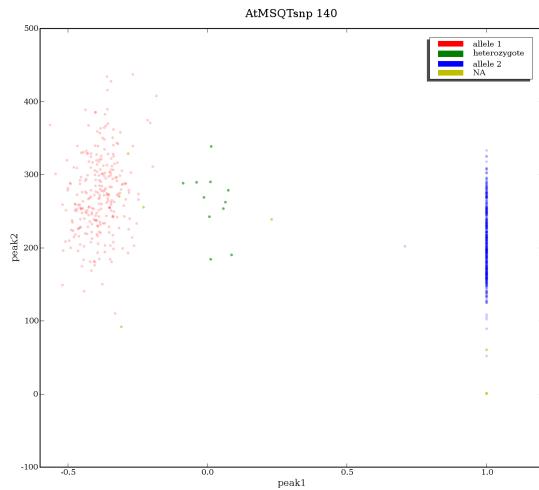


Figure 67: cluster plot for AtMSQTsnp 140.

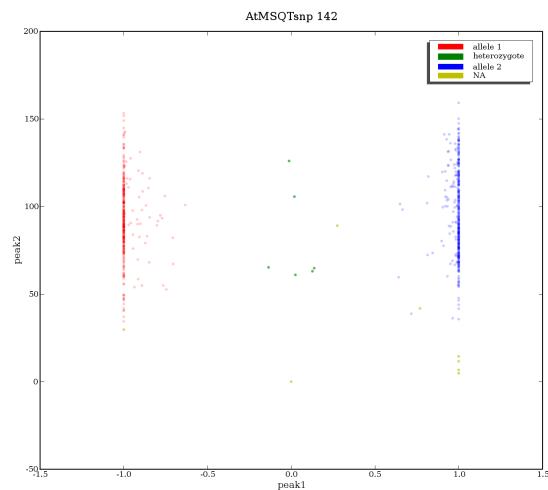


Figure 68: cluster plot for AtMSQTsnp 142.

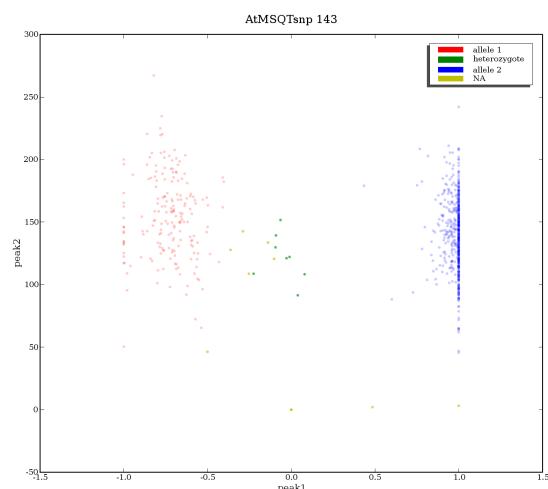


Figure 69: cluster plot for AtMSQTsnp 143.

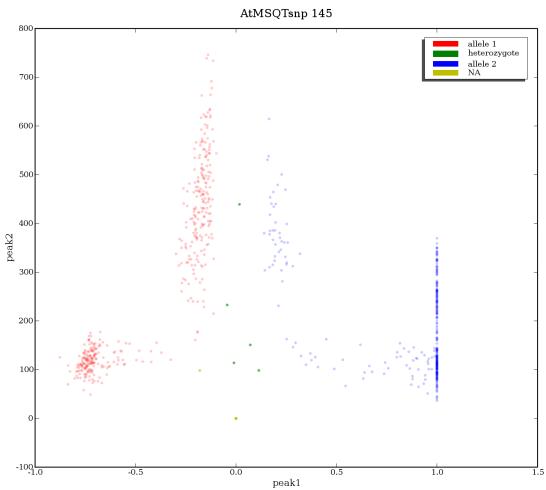


Figure 70: cluster plot for AtMSQTsnp 145.

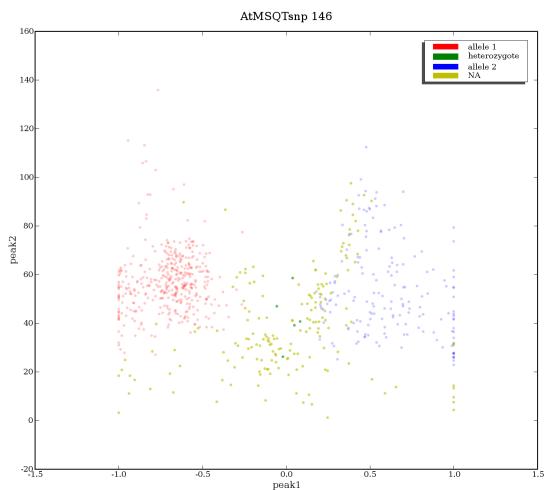


Figure 71: cluster plot for AtMSQTsnp 146.

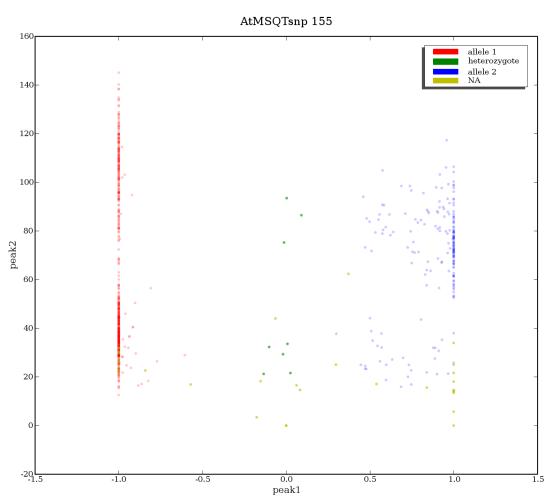


Figure 72: cluster plot for AtMSQTsnp 155.

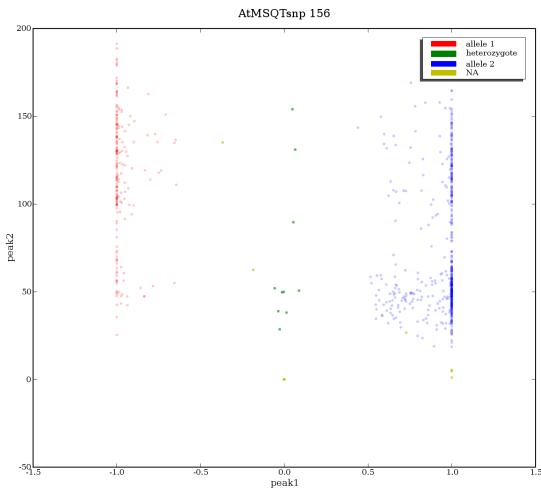


Figure 73: cluster plot for AtMSQTsnp 156.

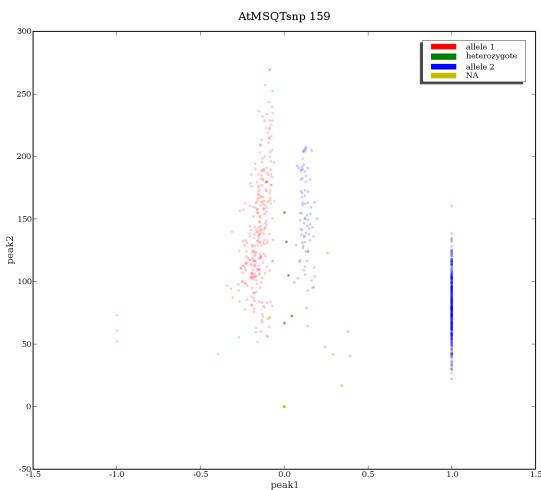


Figure 74: cluster plot for AtMSQTsnp 159.

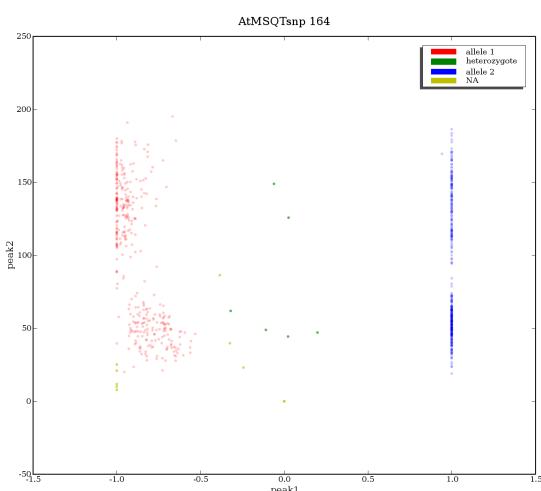


Figure 75: cluster plot for AtMSQTsnp 164.

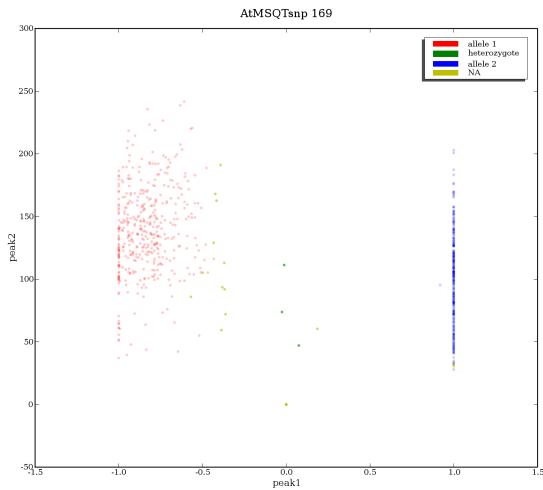


Figure 76: cluster plot for AtMSQTsnp 169.

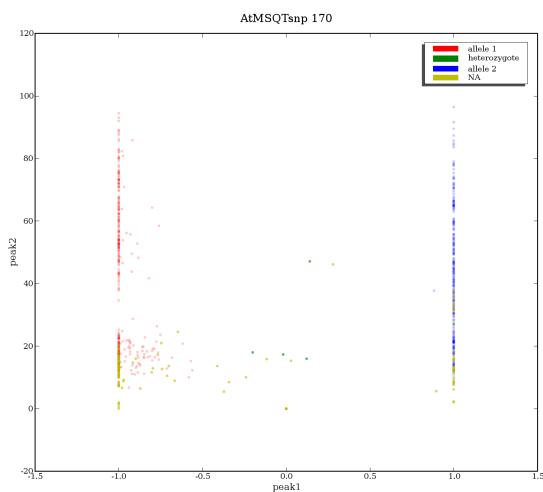


Figure 77: cluster plot for AtMSQTsnp 170.

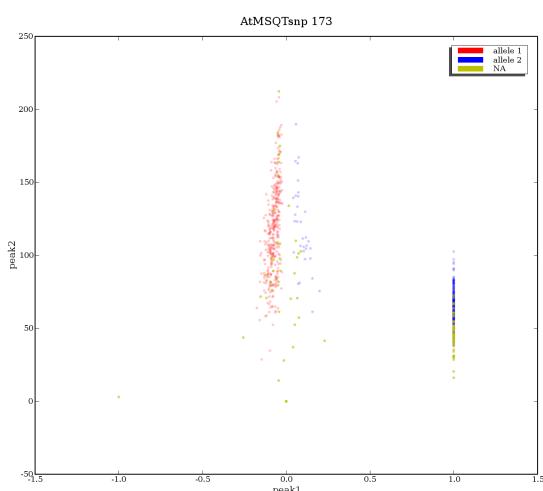


Figure 78: cluster plot for AtMSQTsnp 173.

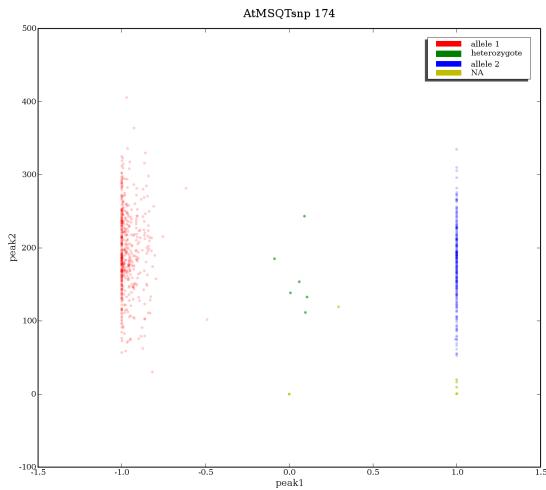


Figure 79: cluster plot for AtMSQTsnp 174.

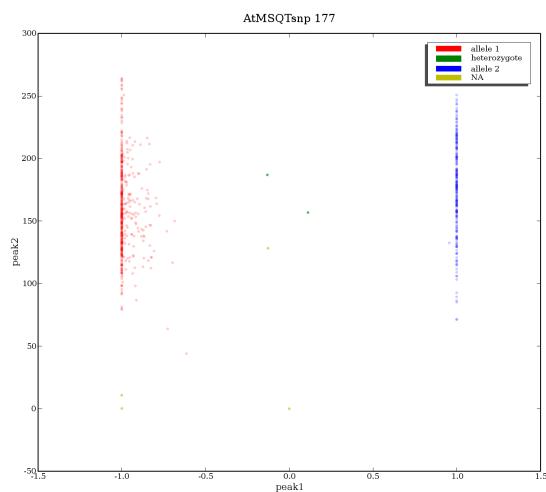


Figure 80: cluster plot for AtMSQTsnp 177.

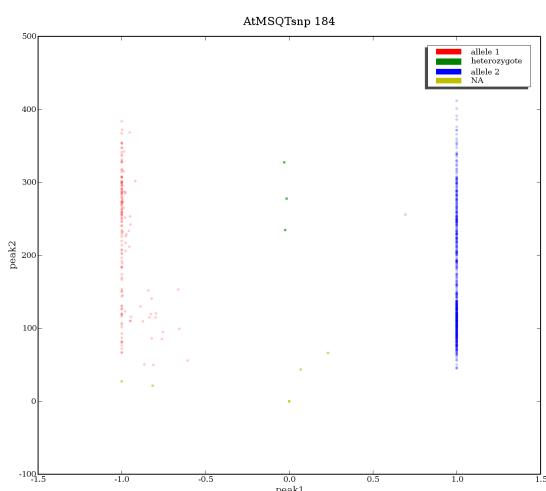


Figure 81: cluster plot for AtMSQTsnp 184.

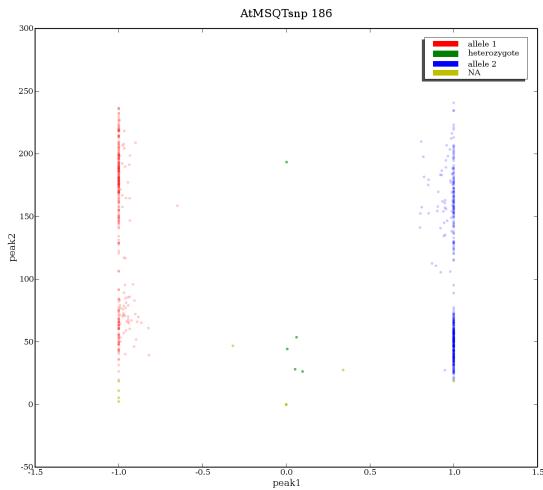


Figure 82: cluster plot for AtMSQTsnp 186.

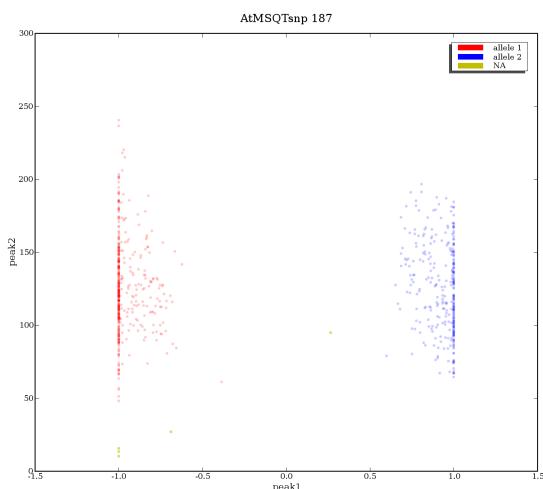


Figure 83: cluster plot for AtMSQTsnp 187.

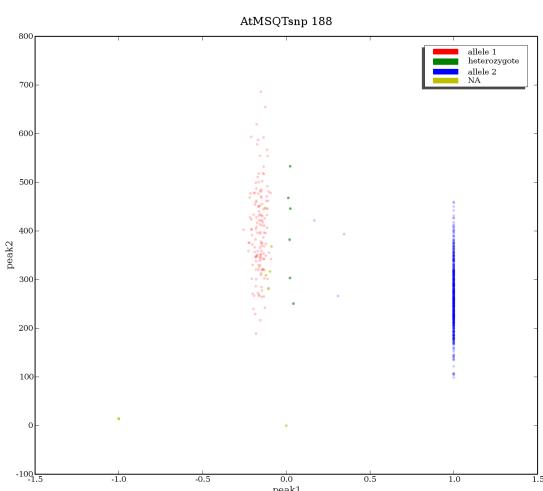


Figure 84: cluster plot for AtMSQTsnp 188.

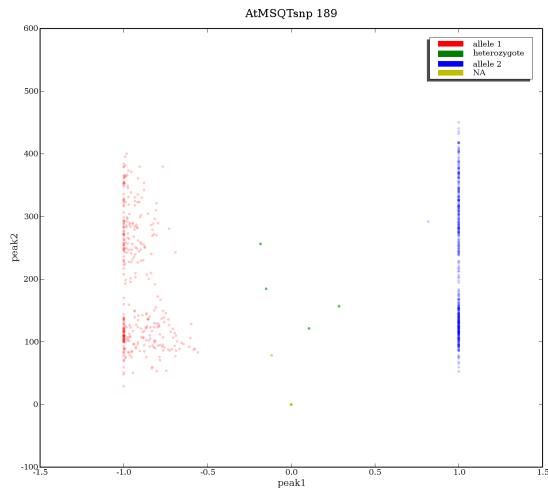


Figure 85: cluster plot for AtMSQTsnp 189.

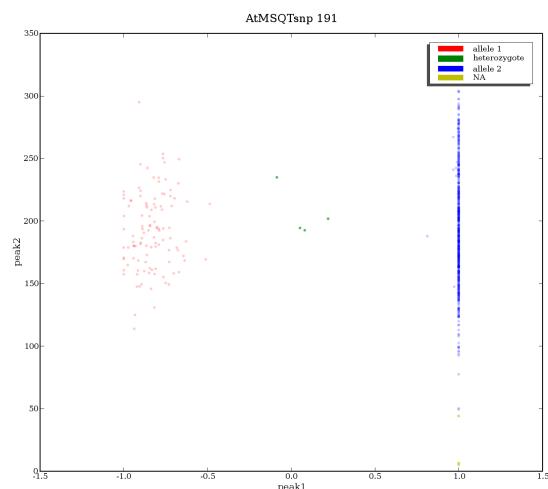


Figure 86: cluster plot for AtMSQTsnp 191.

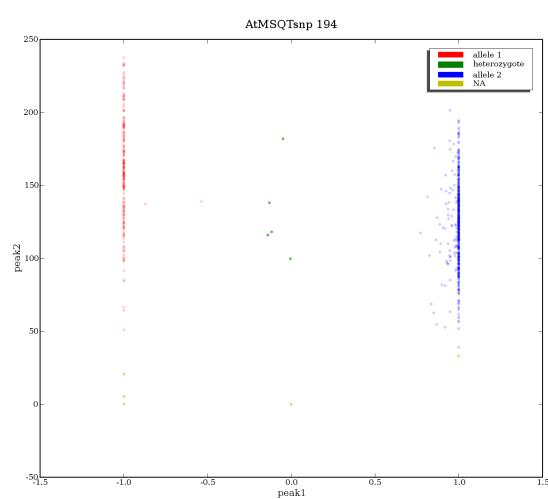


Figure 87: cluster plot for AtMSQTsnp 194.

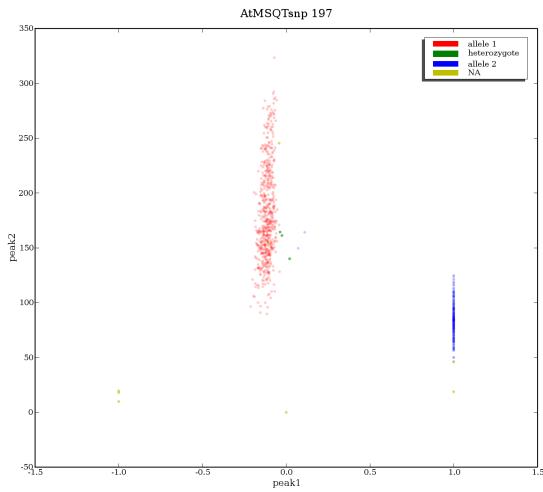


Figure 88: cluster plot for AtMSQTsnp 197.

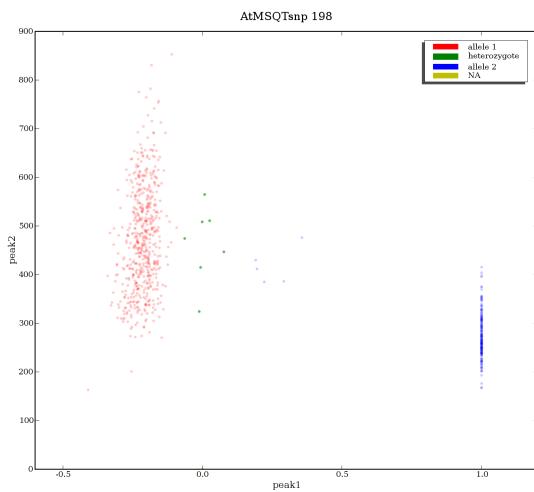


Figure 89: cluster plot for AtMSQTsnp 198.

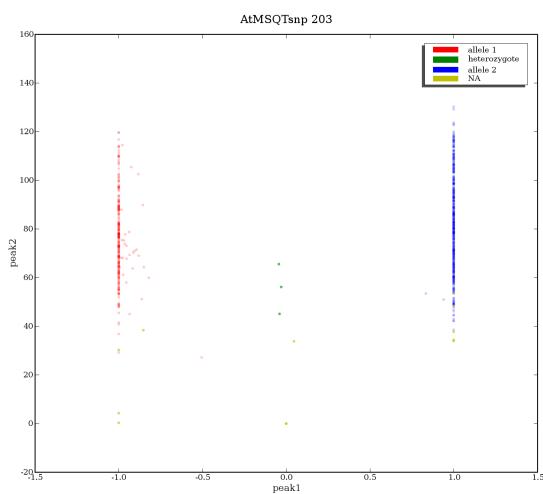


Figure 90: cluster plot for AtMSQTsnp 203.

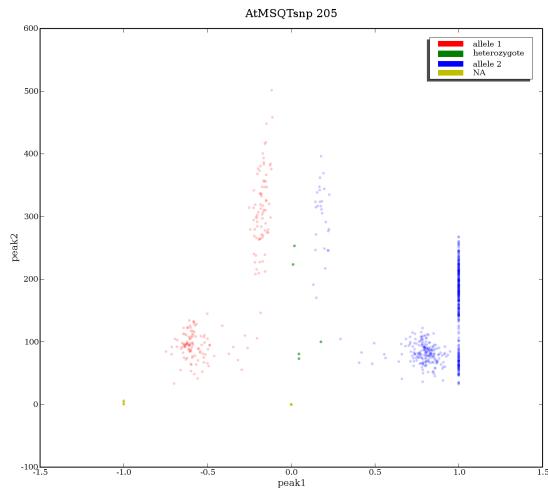


Figure 91: cluster plot for AtMSQTsnp 205.

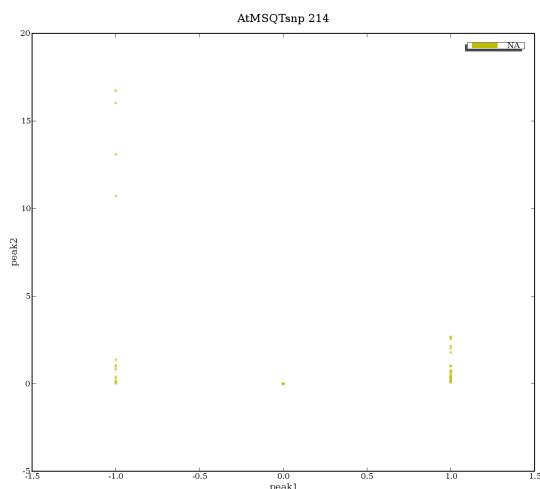


Figure 92: cluster plot for AtMSQTsnp 214.

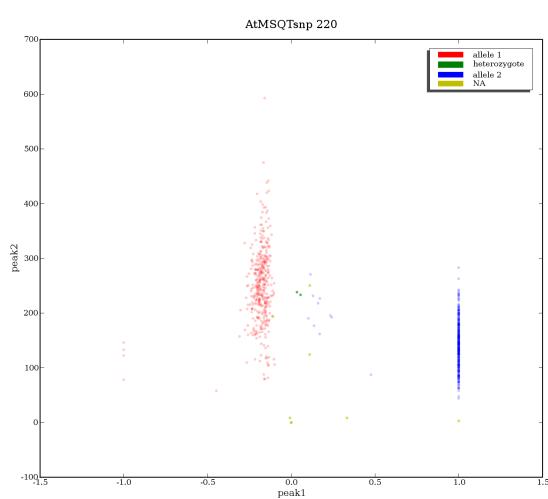


Figure 93: cluster plot for AtMSQTsnp 220.

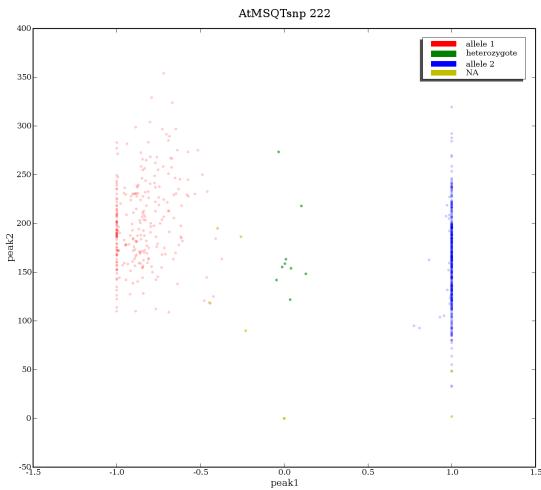


Figure 94: cluster plot for AtMSQTsnp 222.

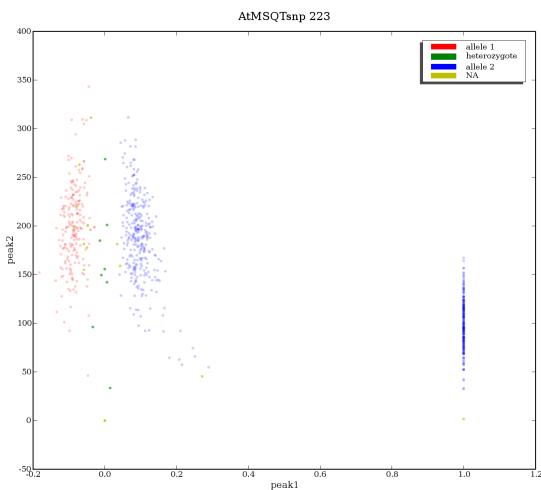


Figure 95: cluster plot for AtMSQTsnp 223.

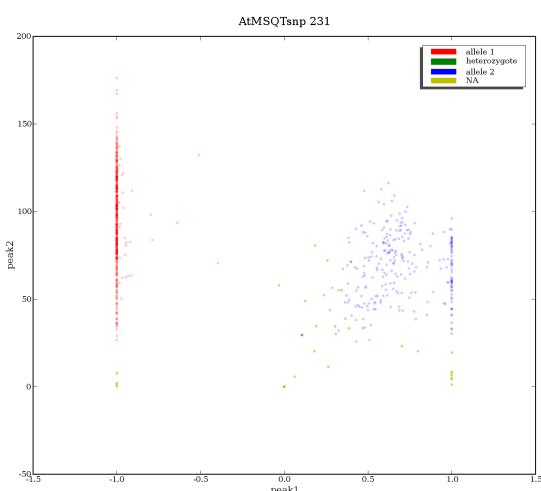


Figure 96: cluster plot for AtMSQTsnp 231.

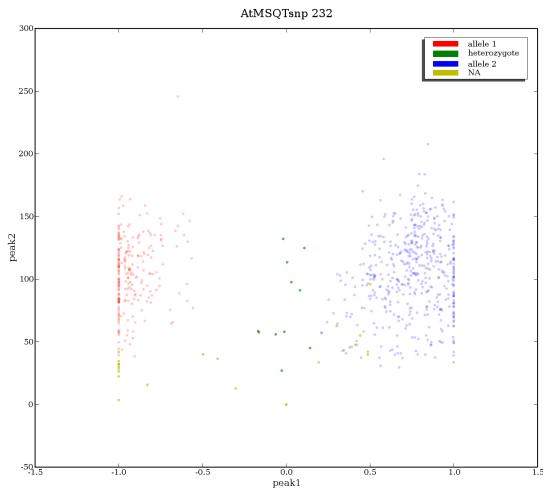


Figure 97: cluster plot for AtMSQTsnp 232.

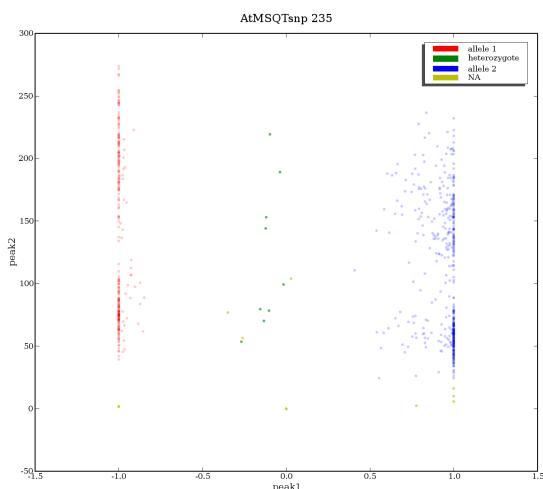


Figure 98: cluster plot for AtMSQTsnp 235.

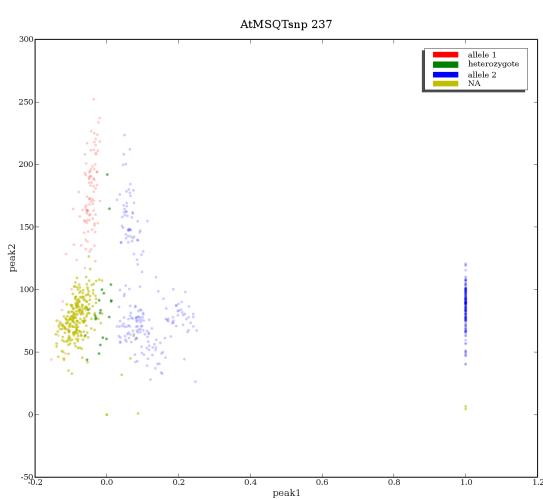


Figure 99: cluster plot for AtMSQTsnp 237.

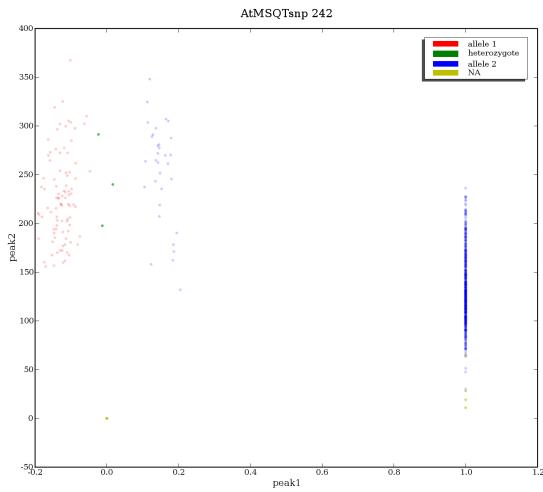


Figure 100: cluster plot for AtMSQTsnp 242.

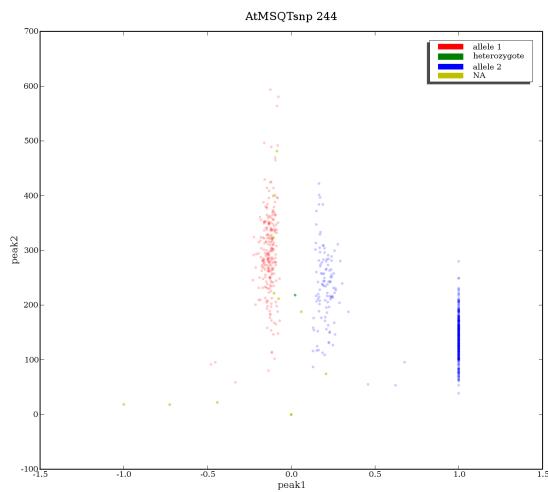


Figure 101: cluster plot for AtMSQTsnp 244.

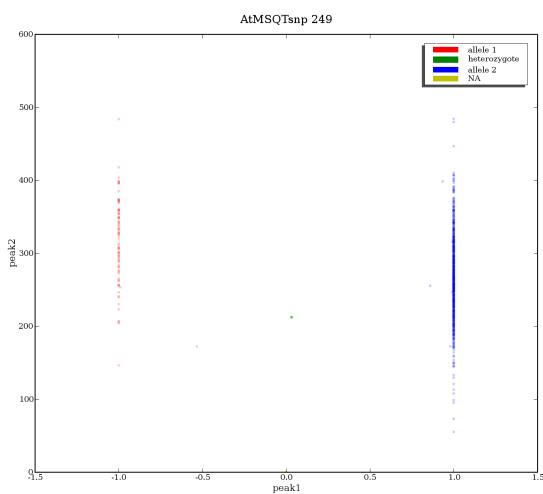


Figure 102: cluster plot for AtMSQTsnp 249.

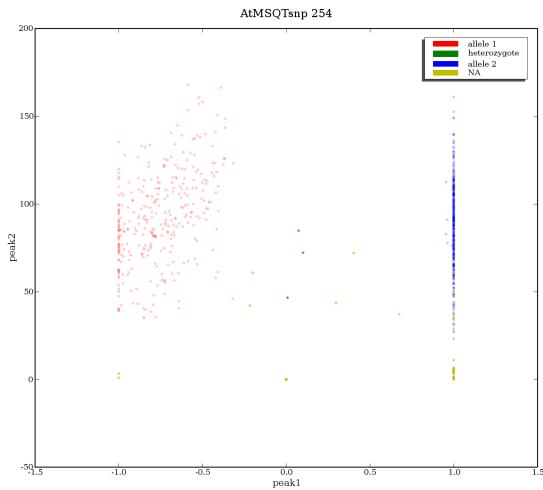


Figure 103: cluster plot for AtMSQTsnp 254.

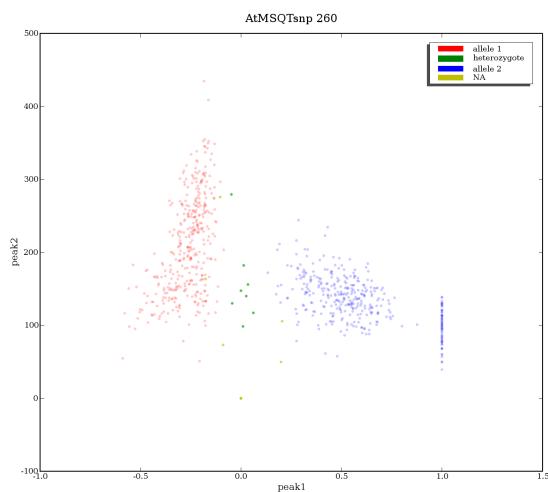


Figure 104: cluster plot for AtMSQTsnp 260.

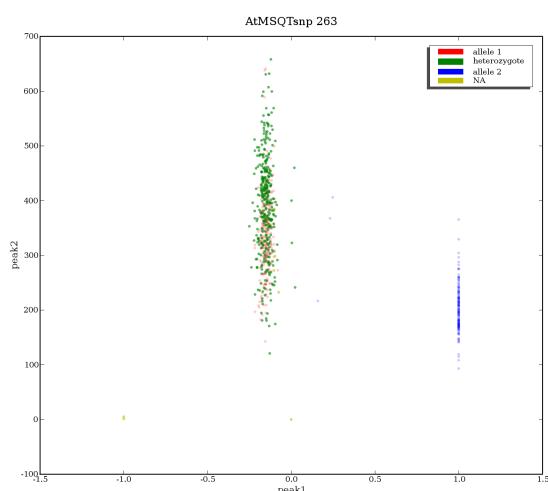


Figure 105: cluster plot for AtMSQTsnp 263.

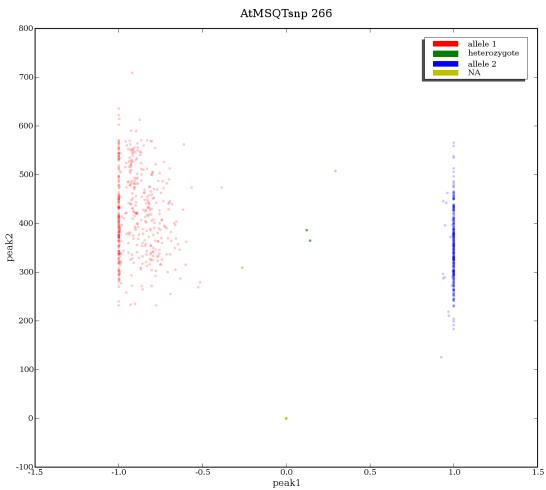


Figure 106: cluster plot for AtMSQTsnp 266.

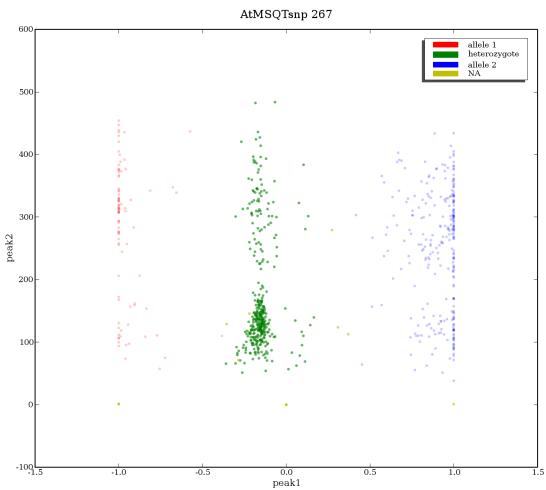


Figure 107: cluster plot for AtMSQTsnp 267.

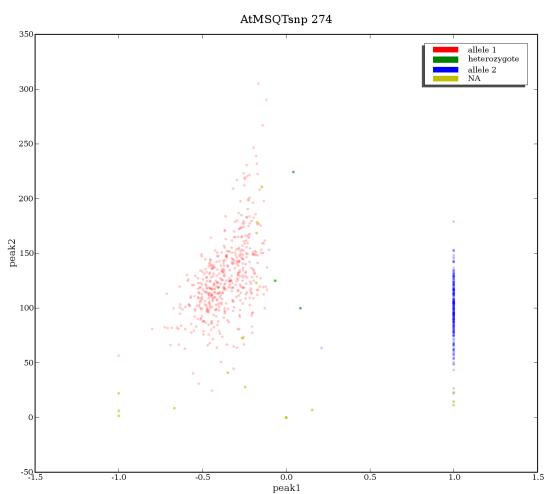


Figure 108: cluster plot for AtMSQTsnp 274.

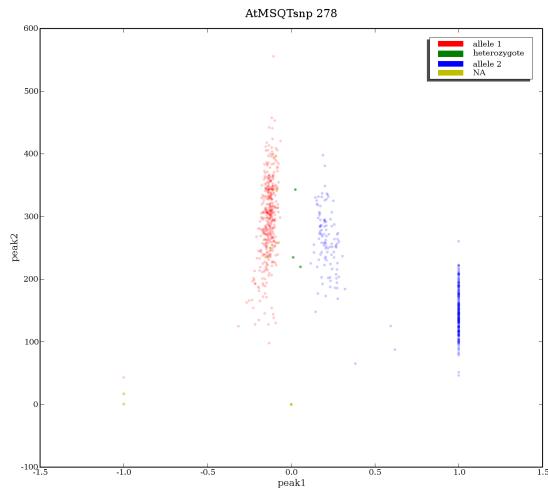


Figure 109: cluster plot for AtMSQTsnp 278.

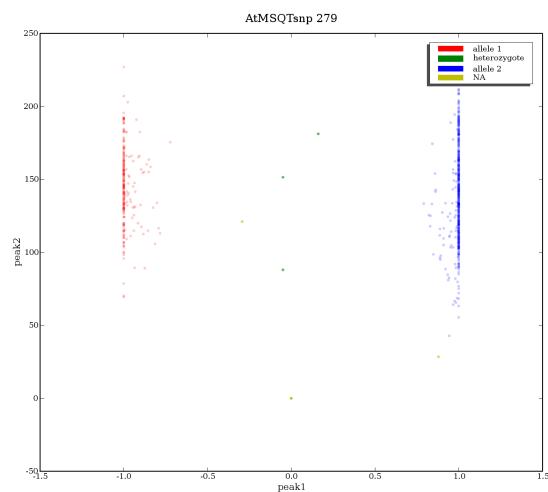


Figure 110: cluster plot for AtMSQTsnp 279.

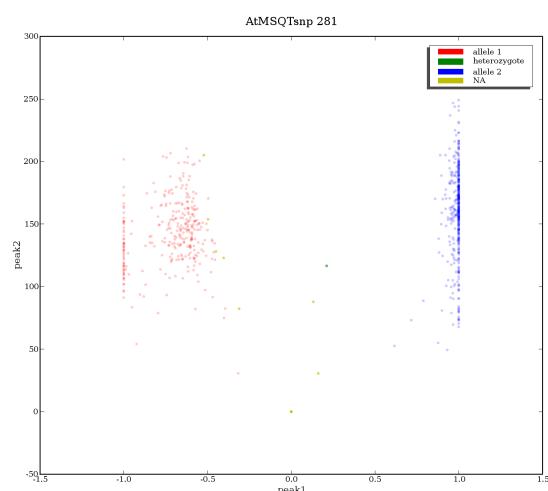


Figure 111: cluster plot for AtMSQTsnp 281.

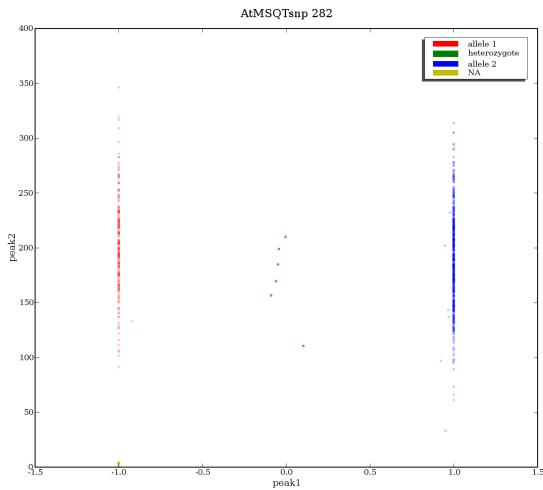


Figure 112: cluster plot for AtMSQTsnp 282.

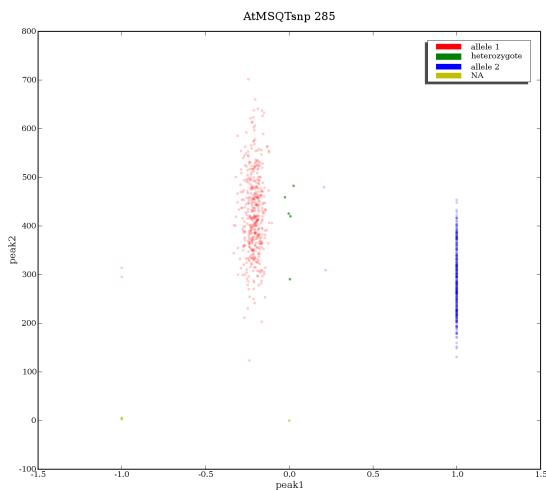


Figure 113: cluster plot for AtMSQTsnp 285.

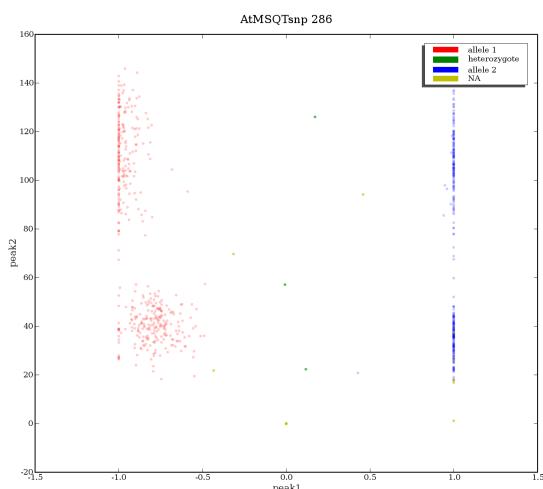


Figure 114: cluster plot for AtMSQTsnp 286.

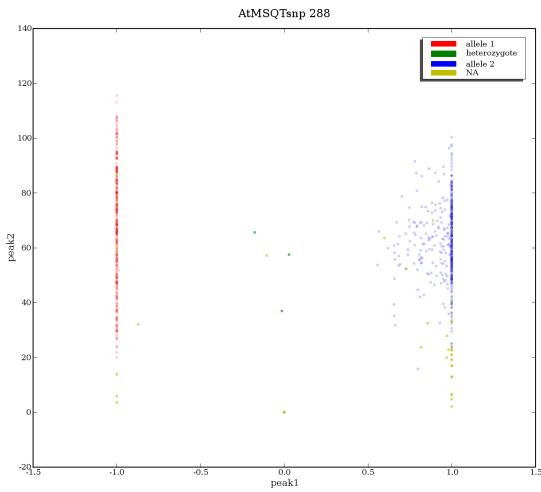


Figure 115: cluster plot for AtMSQTsnp 288.

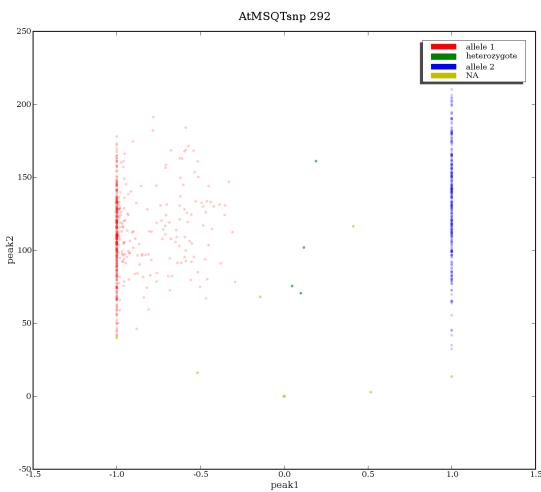


Figure 116: cluster plot for AtMSQTsnp 292.

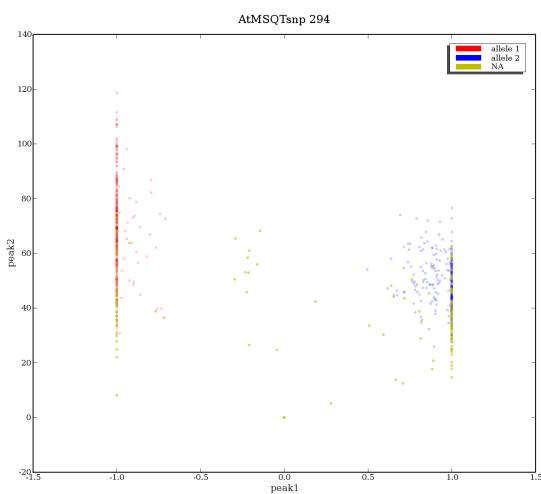


Figure 117: cluster plot for AtMSQTsnp 294.

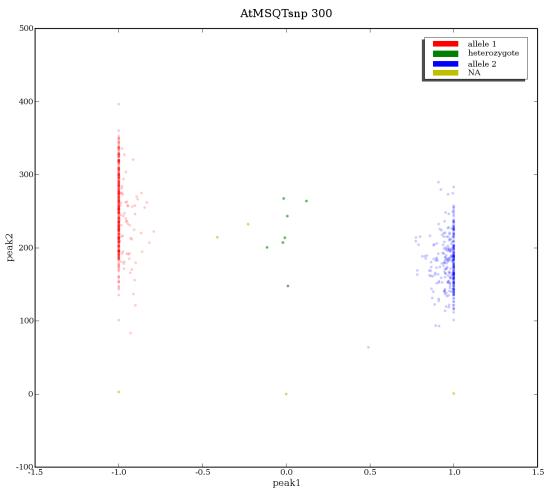


Figure 118: cluster plot for AtMSQTsnp 300.

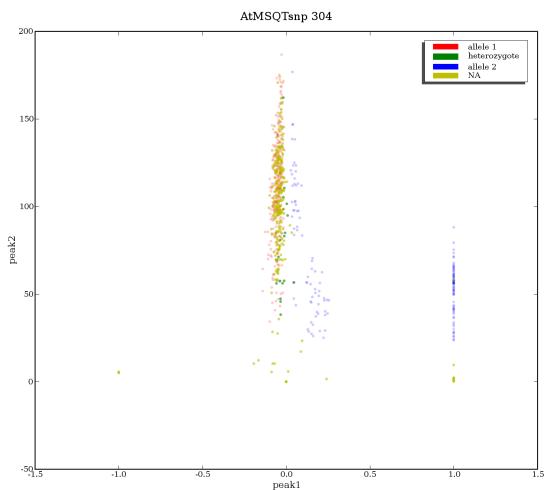


Figure 119: cluster plot for AtMSQTsnp 304.

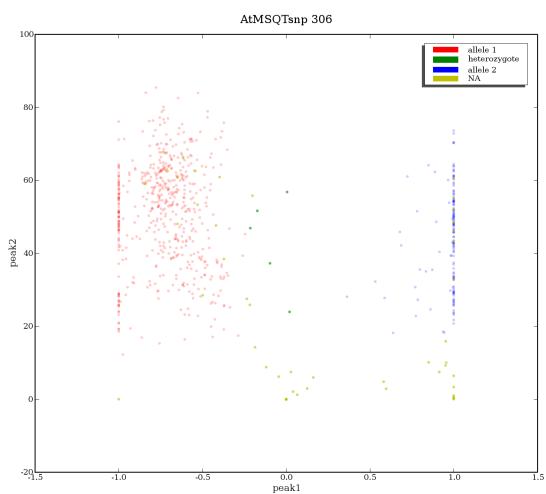


Figure 120: cluster plot for AtMSQTsnp 306.

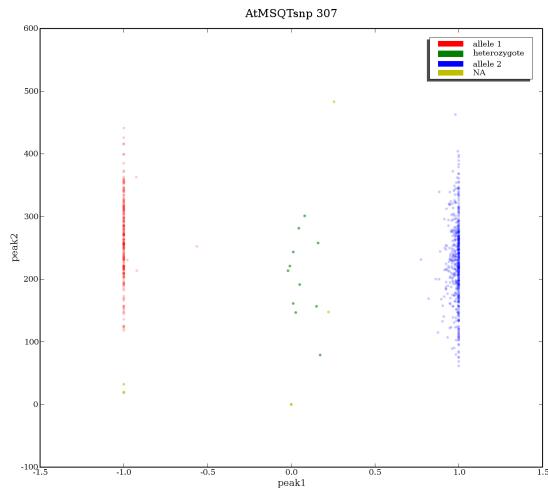


Figure 121: cluster plot for AtMSQTsnp 307.

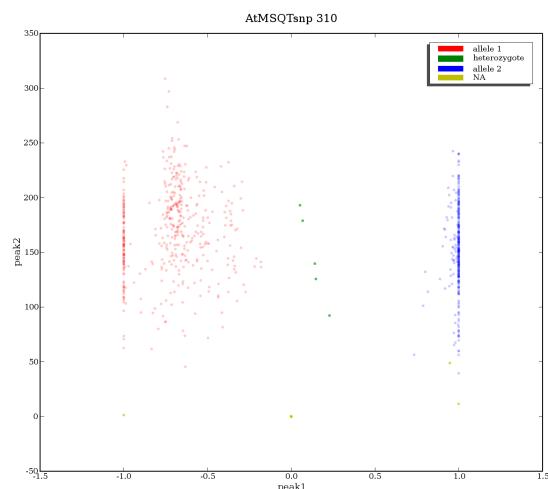


Figure 122: cluster plot for AtMSQTsnp 310.

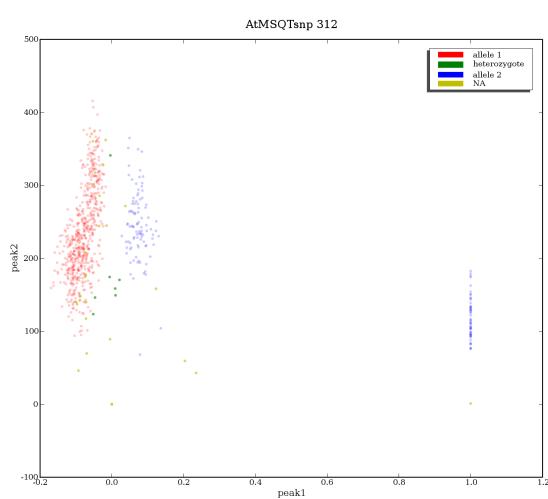


Figure 123: cluster plot for AtMSQTsnp 312.

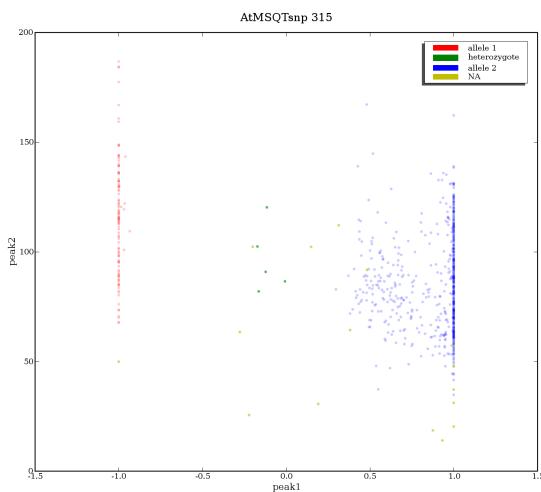


Figure 124: cluster plot for AtMSQTsnp 315.

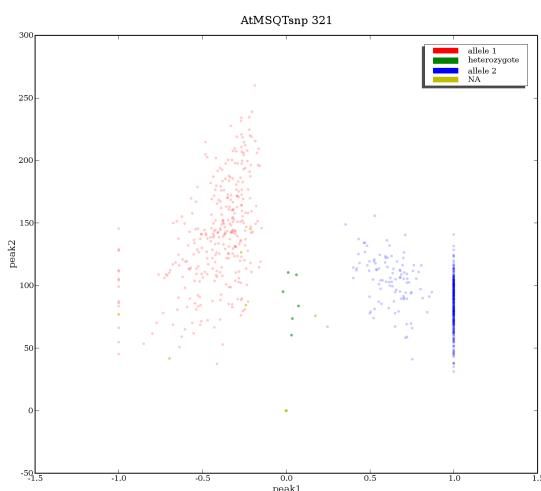


Figure 125: cluster plot for AtMSQTsnp 321.

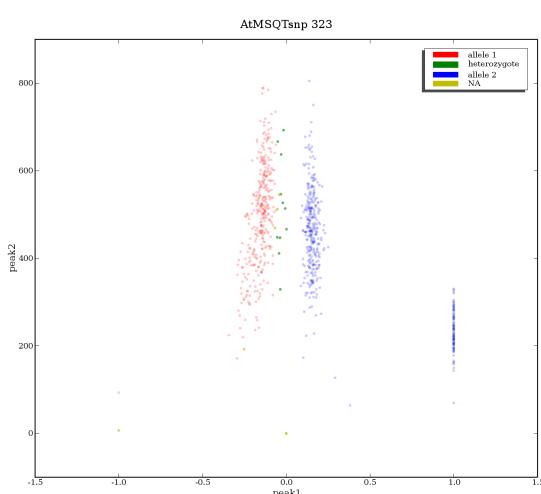


Figure 126: cluster plot for AtMSQTsnp 323.

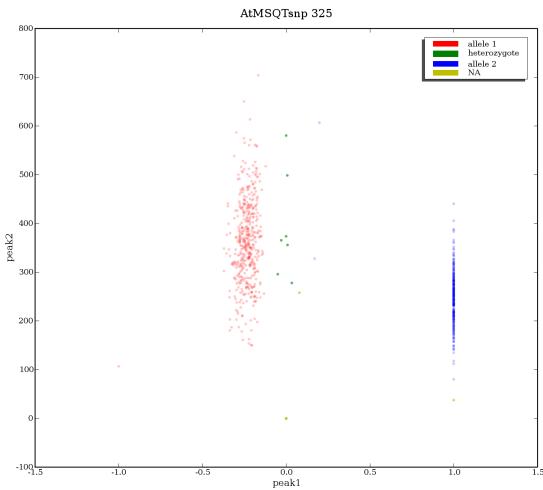


Figure 127: cluster plot for AtMSQTsnp 325.

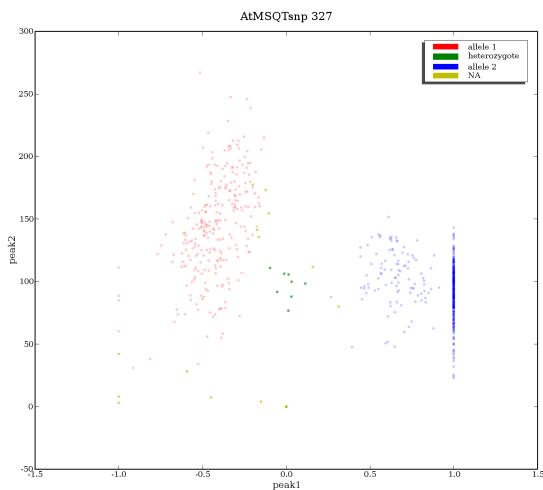


Figure 128: cluster plot for AtMSQTsnp 327.

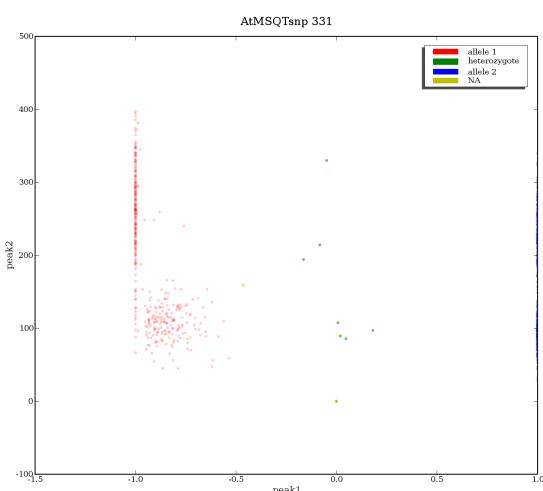


Figure 129: cluster plot for AtMSQTsnp 331.

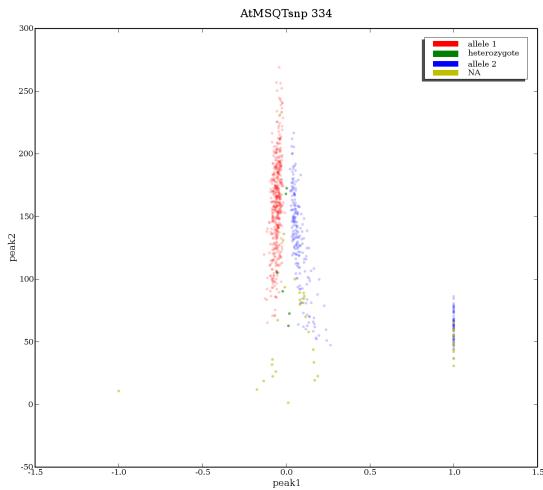


Figure 130: cluster plot for AtMSQTsnp 334.

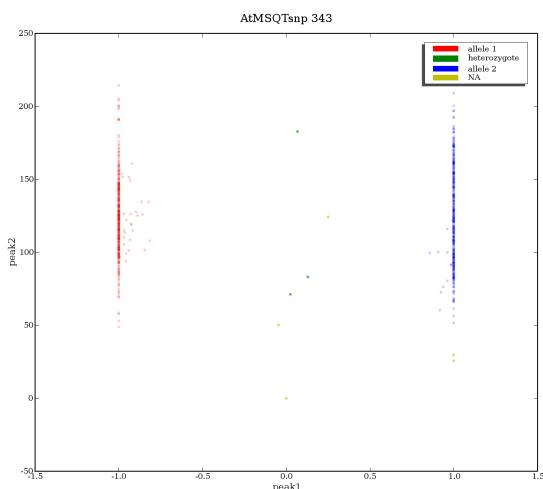


Figure 131: cluster plot for AtMSQTsnp 343.

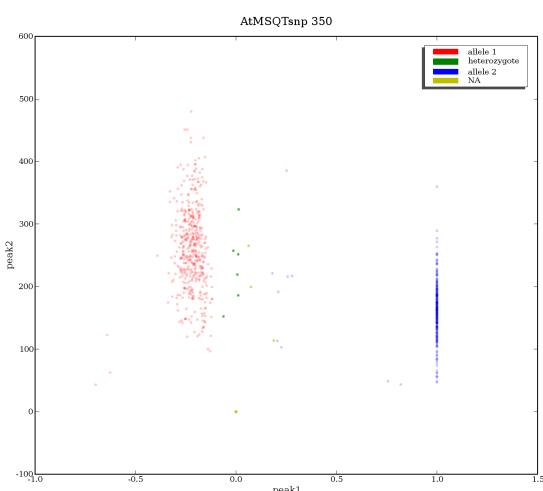


Figure 132: cluster plot for AtMSQTsnp 350.

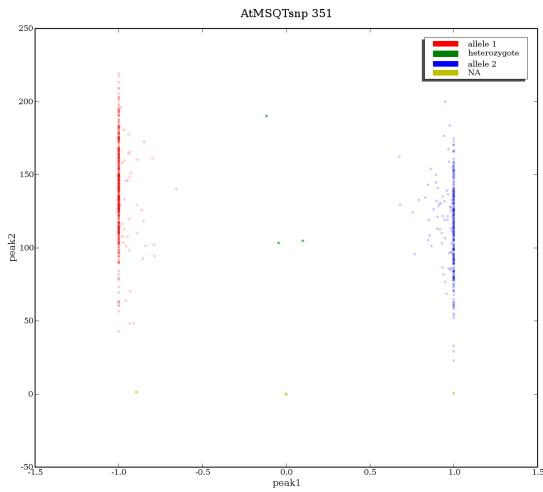


Figure 133: cluster plot for AtMSQTsnp 351.

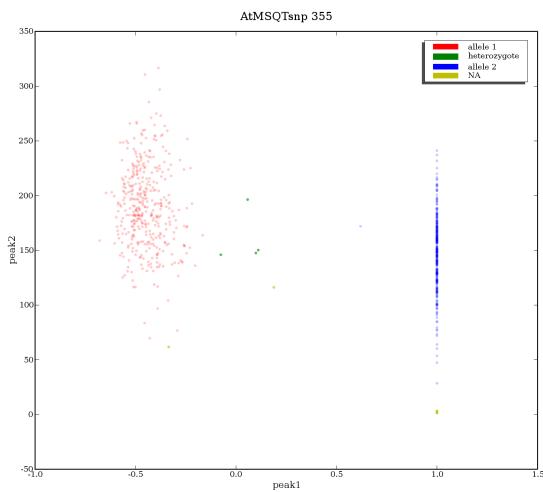


Figure 134: cluster plot for AtMSQTsnp 355.

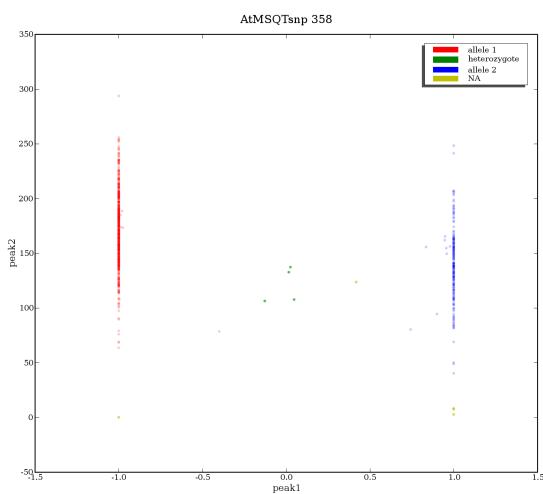


Figure 135: cluster plot for AtMSQTsnp 358.

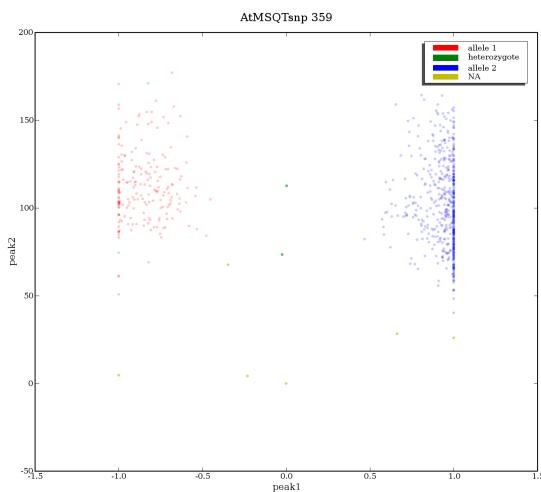


Figure 136: cluster plot for AtMSQTsnp 359.

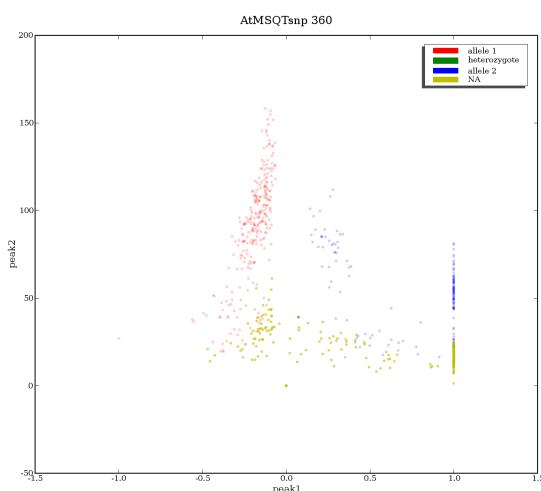


Figure 137: cluster plot for AtMSQTsnp 360.

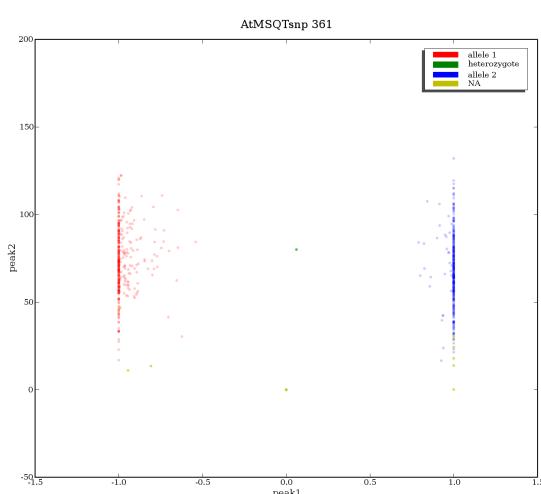


Figure 138: cluster plot for AtMSQTsnp 361.

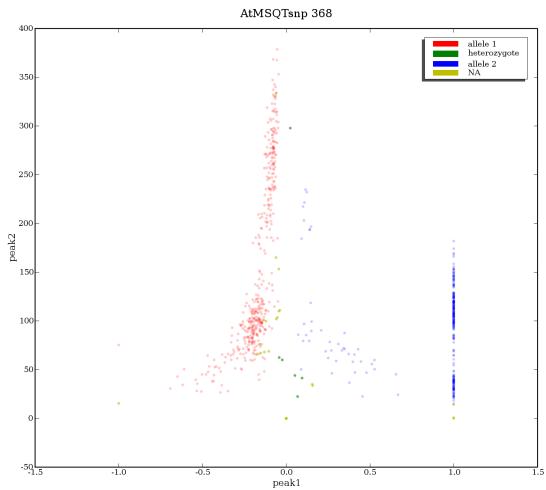


Figure 139: cluster plot for AtMSQTsnp 368.

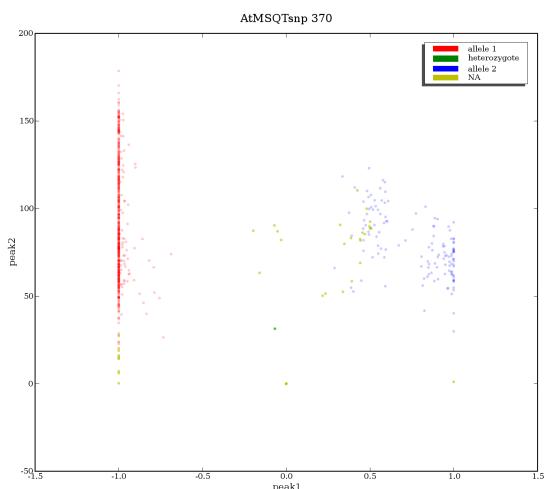


Figure 140: cluster plot for AtMSQTsnp 370.

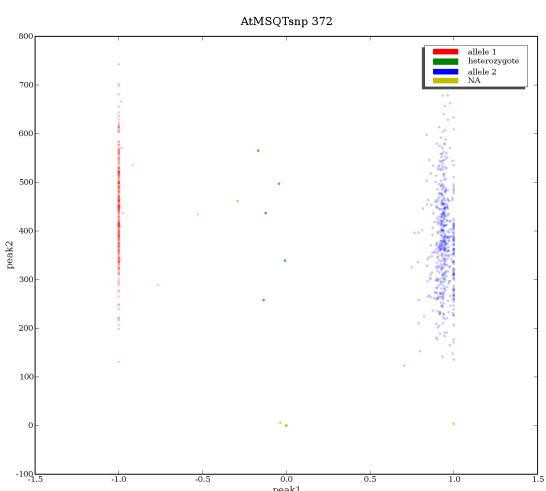


Figure 141: cluster plot for AtMSQTsnp 372.

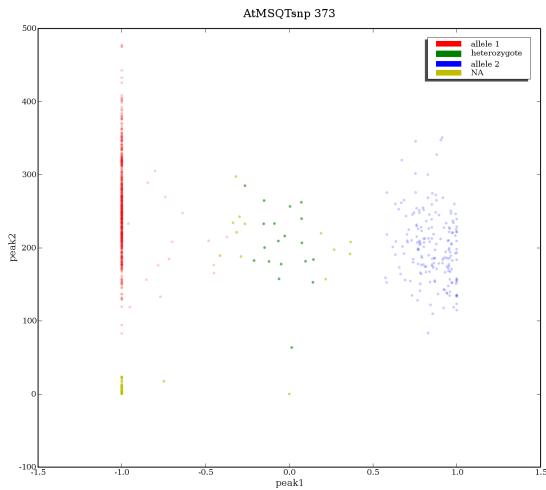


Figure 142: cluster plot for AtMSQTsnp 373.

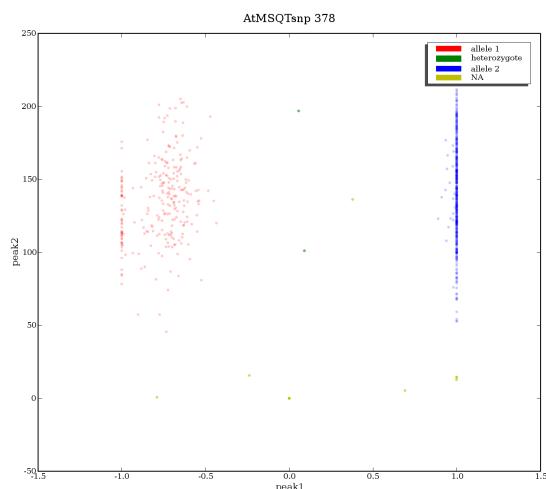


Figure 143: cluster plot for AtMSQTsnp 378.

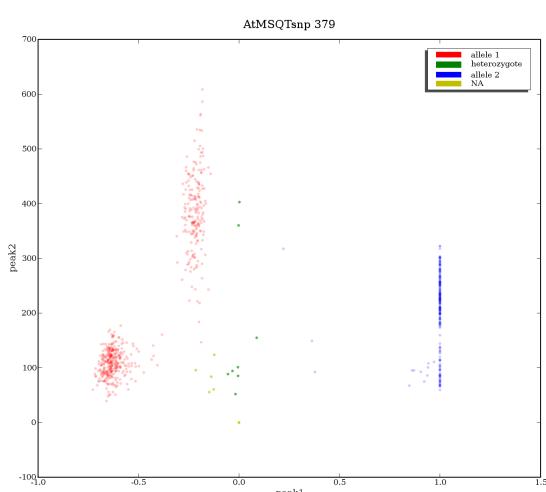


Figure 144: cluster plot for AtMSQTsnp 379.

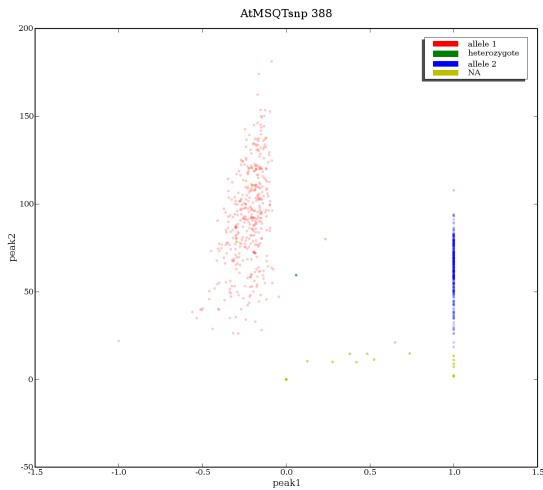


Figure 145: cluster plot for AtMSQTsnp 388.

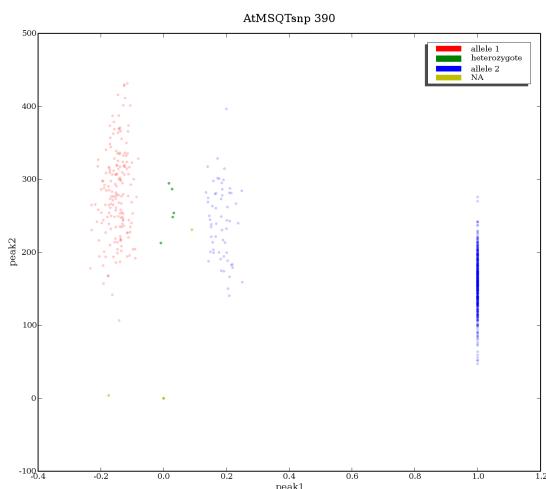


Figure 146: cluster plot for AtMSQTsnp 390.

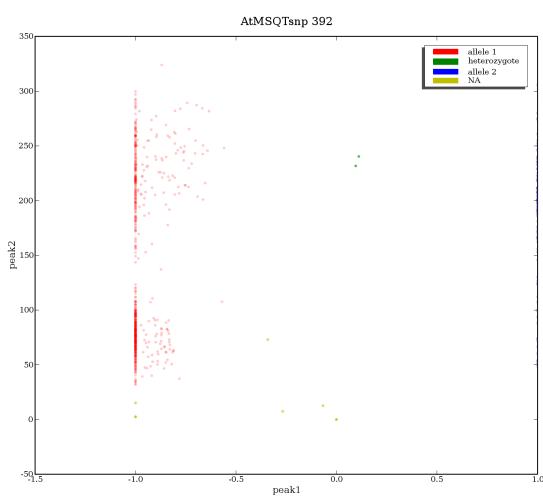


Figure 147: cluster plot for AtMSQTsnp 392.

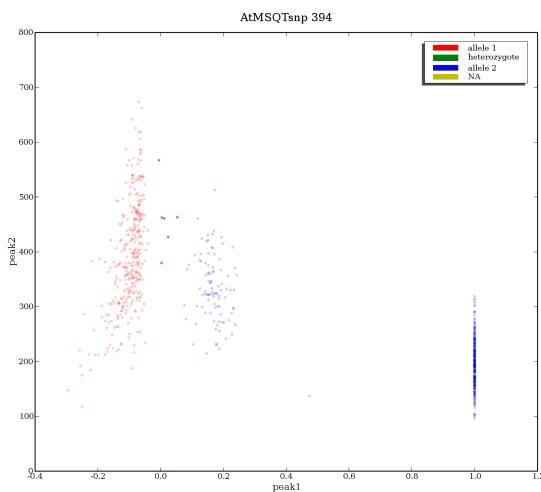


Figure 148: cluster plot for AtMSQTsnp 394.

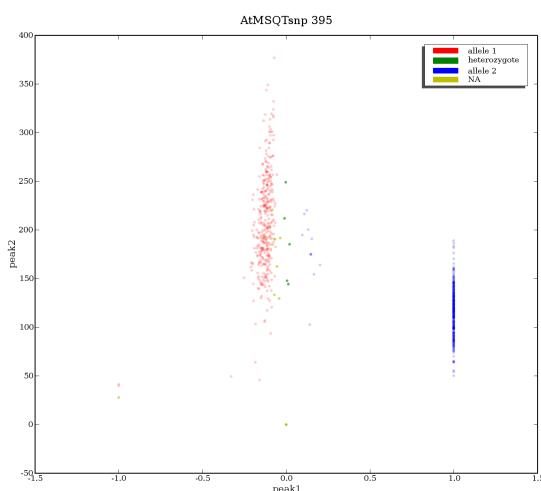


Figure 149: cluster plot for AtMSQTsnp 395.

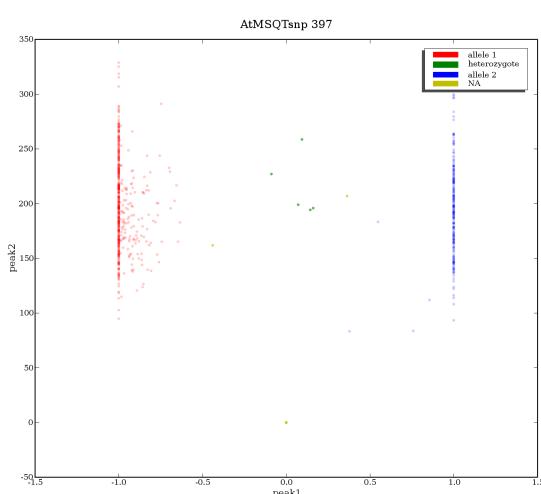


Figure 150: cluster plot for AtMSQTsnp 397.

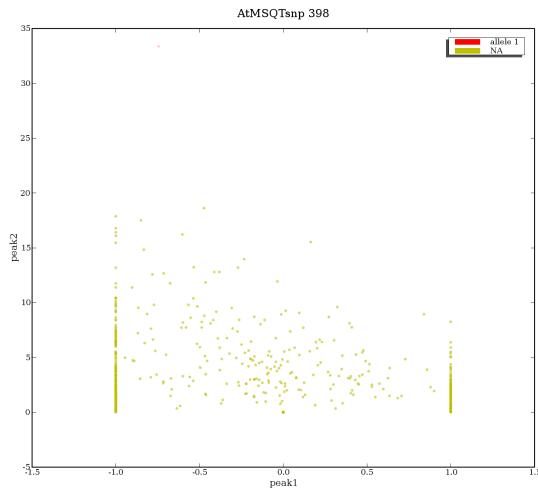


Figure 151: cluster plot for AtMSQTsnp 398.

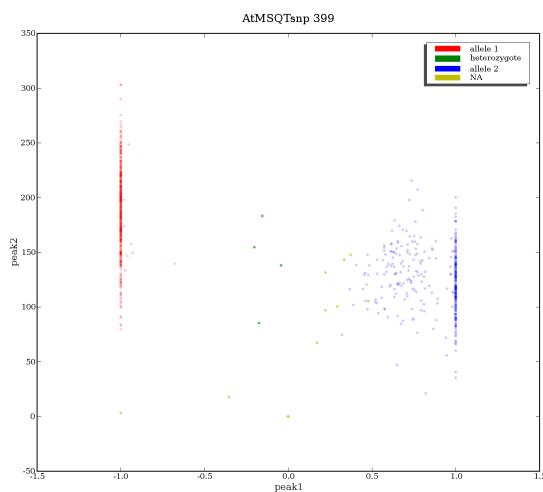


Figure 152: cluster plot for AtMSQTsnp 399.

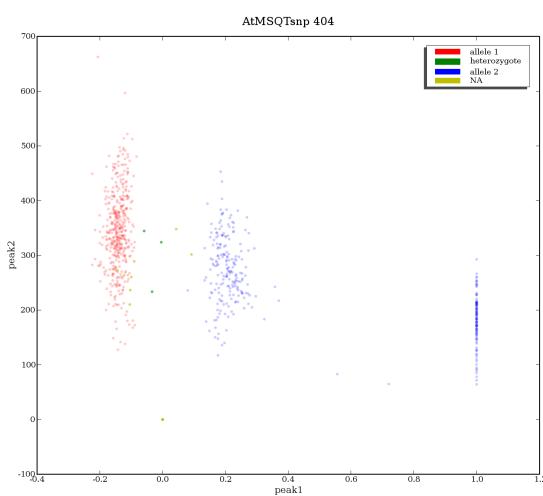


Figure 153: cluster plot for AtMSQTsnp 404.

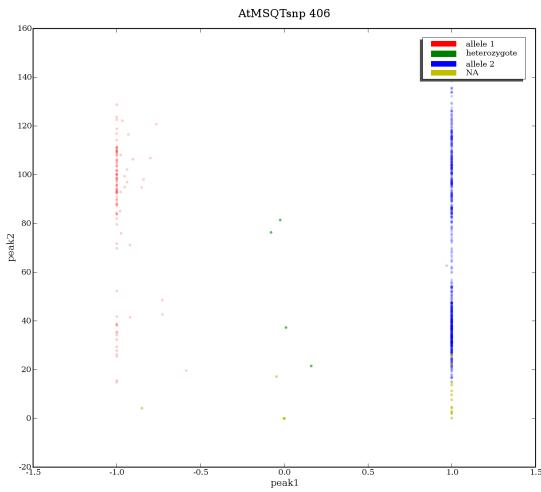


Figure 154: cluster plot for AtMSQTsnp 406.

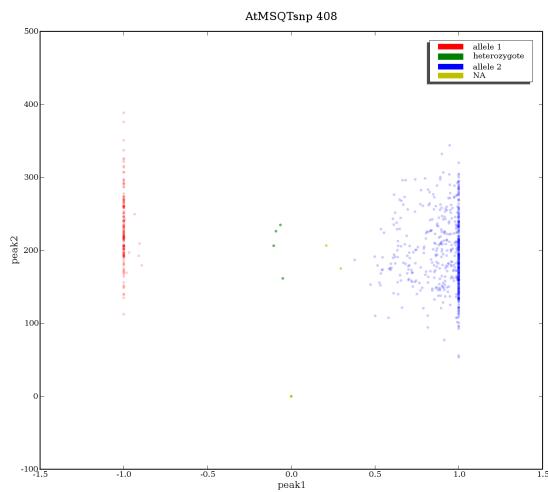


Figure 155: cluster plot for AtMSQTsnp 408.

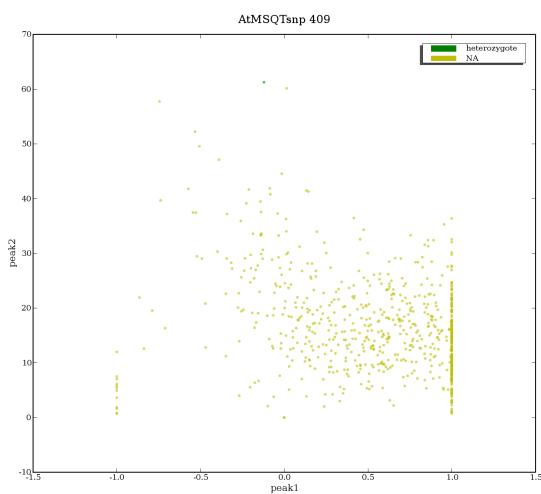


Figure 156: cluster plot for AtMSQTsnp 409.

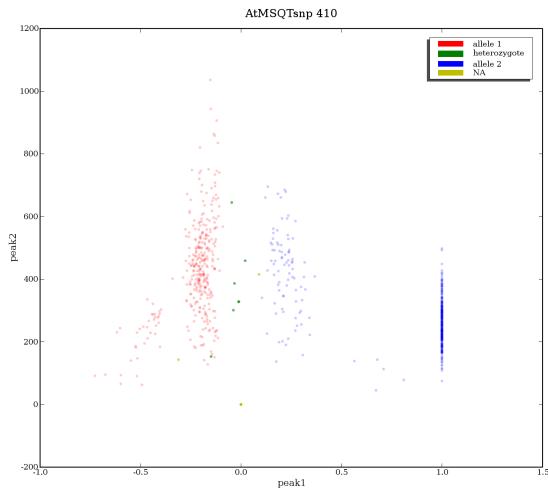


Figure 157: cluster plot for AtMSQTsnp 410.

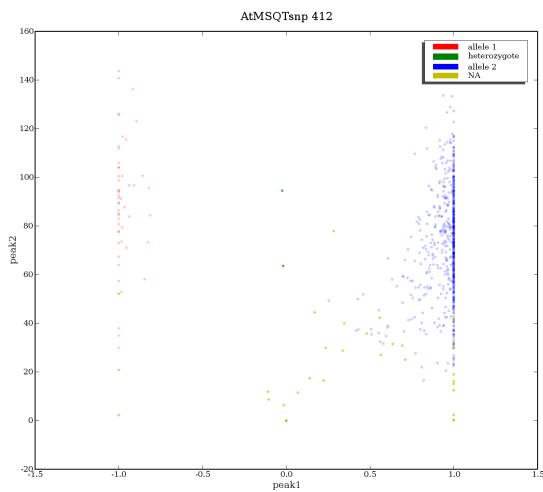


Figure 158: cluster plot for AtMSQTsnp 412.

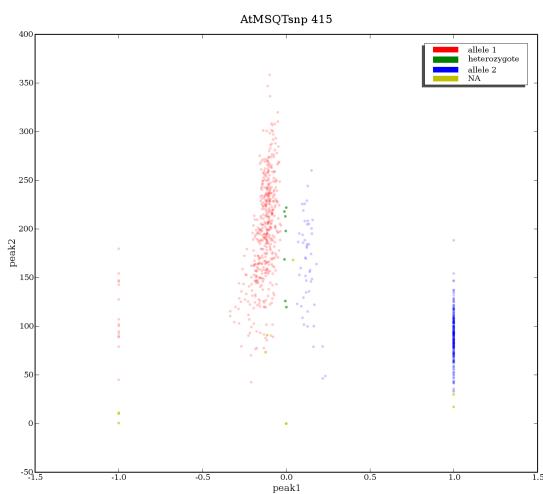


Figure 159: cluster plot for AtMSQTsnp 415.