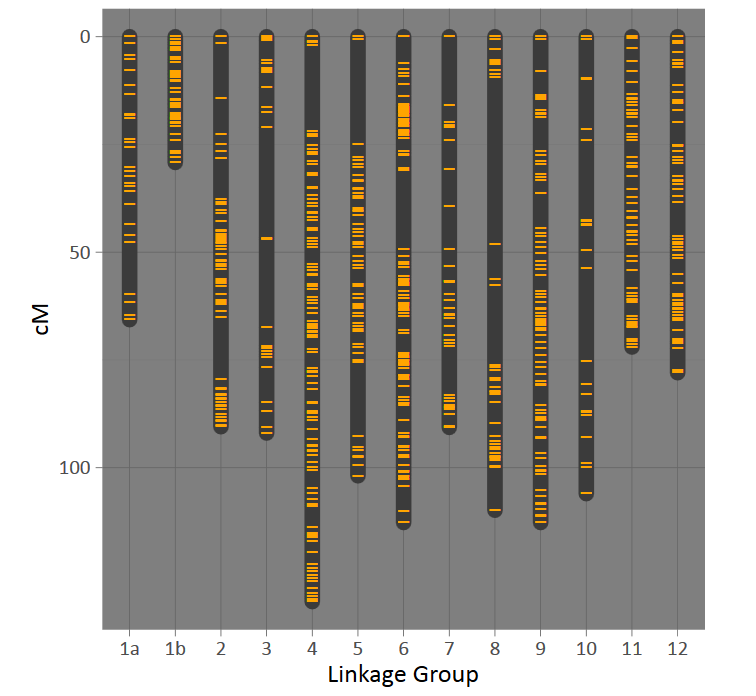
**Mogelijke vragen bij deze dataset**

**Q:**

1. **Chromosome 8 has a fairly large gap of 45 cM in the upper region. At what distance in centi Morgans are markers considered completely unlinked and why?**
2. **Inspect the scatter plot comparing the physical distance in bp with the genetic distance for chromosome 8. Are the markers correctly ordered according to the genome? What is your opinion on ordering markers based on genome assembly? (Hint: genomic rearrangements, paper of Salzberg 2005 – bioinformatics)**
3. **What are implications for genetic maps with large gaps for QTL mapping?**



*Answer:*

*a) At 50 cM, 50% chance of recombination*

*b) They are correctly ordered according to the physical genome assembly. However, assemblies are not 100% accurate or this specific population could have a genomic rearrangement. Ergo, combining a genetic map based on prior knowledge of genome can be dangerous.*

*c) Ghost QTL may appear and imputation of genotypes in QTL mapping can be highly erroneous.*

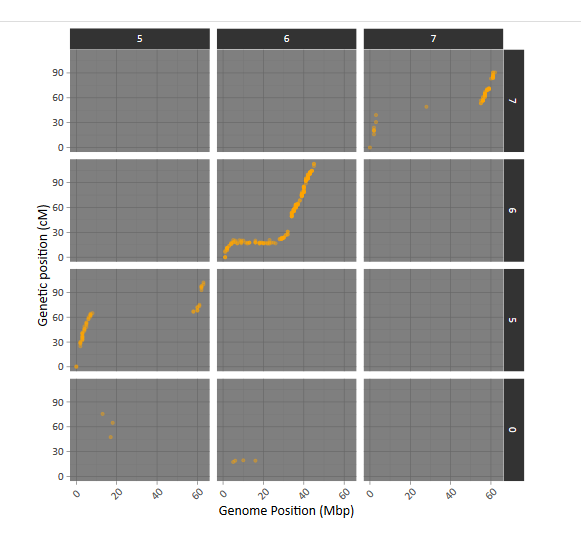
**Q: For one individual there is quite a bit of missing information. What could be possible reasons?**



*Answer: Poor quality DNA due to DNA isolation. Low sequence coverage due to chance or due to non-*equimolar distribution of sample DNA before that is pooled before sequencing

**Q:**

1. **The comparison of physical distance and genetic distance shows a non-linear relationship, why?**
2. **Overall there are relatively few markers present in the middle of the chromosome. What is a possible reason? (Hint: with which technology was genotyping performed in this population?)**

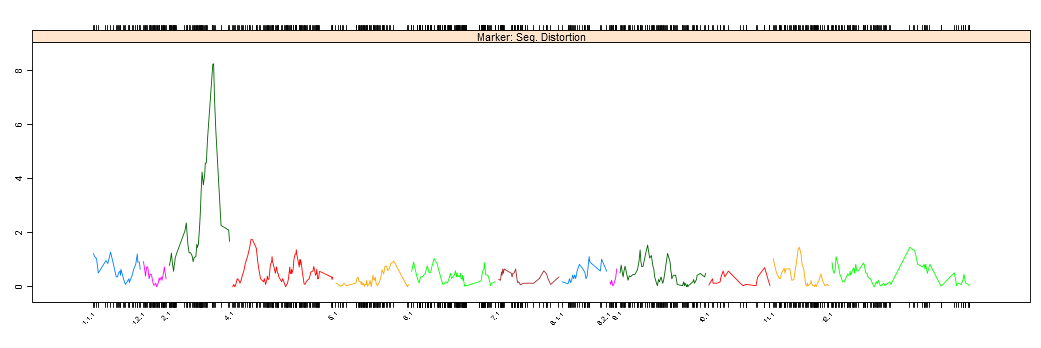


*Answer:*

1. The physical distance of the centromeric region is considerably in tomato. DNA in the centromeres is tightly packed in heterochromatin containing large amounts of silenced repetitive DNA elements. Consequently few recombination can occur in this densely packed genomic region.
2. The centromeric region contains a large amount of repetitive DNA, which is difficult to align. The authors used DNA sequencing for genotyping hence the reads mapping to the centromere could not be mapped with equal reliability compared to telomeric regions. This results in low marker density in the centromeric regions.

**Q:**

**On chromosome 2 there is a region with segregation distorted marker, which means that the distribution of “a” and “b” alleles deviates from the expected 50-50 ratio. What is a likely explanation in this instance: a technical issue or a biological issue with the data? Explain**



*Answer: Likely biological because the distortion shows decay in linkage and multiple marker show segregation distortion. There must have been selection against this locus during populatiin construction, e.g. due to poor germination.*