**Comparative Analysis Assignment**

**General statistics:**

1. **Which genomes did you choose?**

I choose *Brassica oleracea var. capitata* (Golden Early) (v1.1, id24777) and *Brassica rapa* (v1.5, id24668).

1. **What was your rationale?**

They diverged from a common ancestor and are a part of the Triangle of U, hence I found these quite interesting. Also, I love the aroma of flowers of *B.rapa* and I love cabbage with peas (my mom used to cook an amazing dish for me using these two).

1. **Are these two genomes well assembled?**

Yes, I think so. As there is only 4.96% means about 5% of the noncoding sequence is uncertain for B. rapa and B. oleraceae N is 6.66%.

1. **How many contigs/scaffolds are there?**

B. rapa have 40,367 and B. oleraceae have 9 contigs.

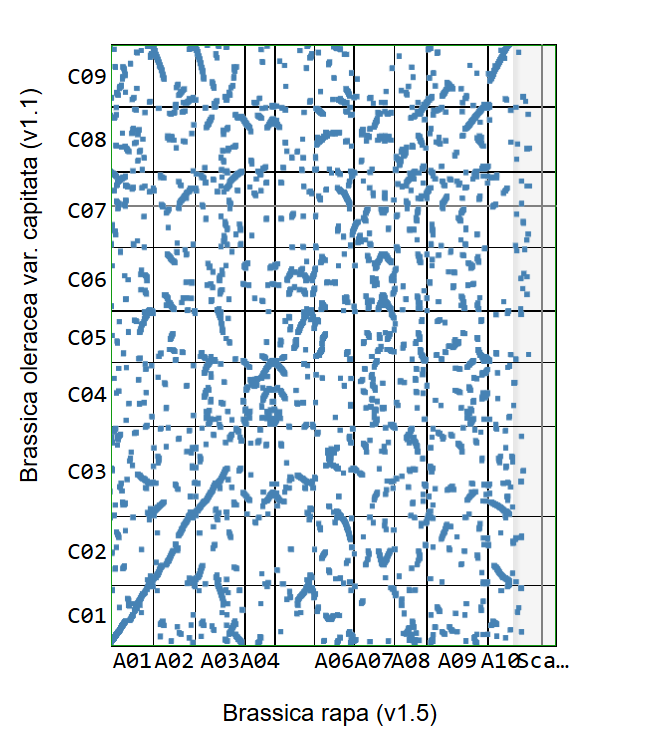
1. **Are your species diploid, polyploid, or a combination of both?**

Both are diploid in nature but B. oleracea have chromosome number: 2n = 2x = 18 and B.rapa have chromosome number: 2n = 2x = 20.

1. **Can you tell who sequeunced that genome or what technology was used (the answer might be nFrom the file, we can see these reference accessions:**

From the file, I can see Brassica oleracea var. capitata (Golden Early) (v1.1, id24777) was sequenced by Feng Cheng and Brassica rapa (v1.5, id24668) by Eric Lyons.

**Macrosynteny**

****

1. **Do these two species/ or genome versions have good macrosynteny? Is there clear evidence of polyploidy?**

Yes, these two genomes have a strong macrosynteny. There are long, diagonal blocks indicating large collinear segments between the A and C genomes, consistent with the close relationship of *B. rapa* (A genome) and *B. oleracea* (C genome). As for the evidence of polyploidy, individual C-chromosome segments in B. oleracea map to multiple blocks across B. rapa A-chromosomes (and vice-versa), replicated, near-parallel blocks rather than a single one-to-one diagonal.

1. **How similar do you think these species are? Do the syntenic dotplots match your expectations based on how close they are phylogenetically?**

*B. rapa* and *B. oleracea* are sister Brassica crops; the extensive collinearity and multi-block correspondences match expectations for very closely related species that diverged after the shared ancestor.

1. **Are there any inversions or large-scale chromosomal rearangements? If so, list a few (e.g., Sorghum choromosome 1 and Rice chromosome 7 are syntenic, but there is a large inversion) Note: this might be very difficult to assess if you picked highly divergent species.**

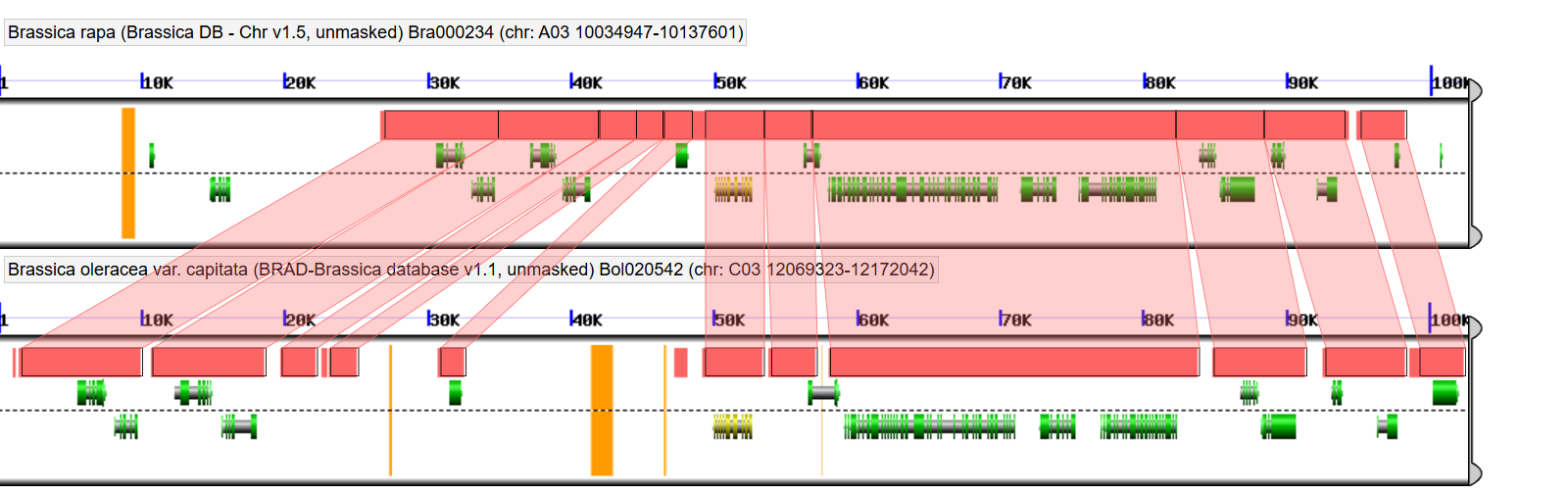
Yes, there are segments that align with opposite orientation relative to their partners which imply inversions and chromosomal rearrangements during diploidization. For instance, at A03’C02, A06’C02, and inversion at A04’C04 highlights local orientation flips within an otherwise syntenic block.

1. **Are these genomes well assembled or is one/both highly fragmented? How does that affect your interpretation?**

The *B. oleracea* assembly is chromosome-level (9 chromosomes listed), whereas your *B. rapa* selection is flagged as “Contains contigs,” so it’s more fragmented. Fragmentation created broken diagonals and small off-axis blocks that show rearrangements. Despite that, the major cross-chromosome diagonals are still very clear, so assembly fragmentation doesn’t obscure the main conclusions.

**Microsynteny**

**Select one region and look closer at it using GEvo. Paste a screenshot of this region.**

****

1. **What chromosomal regions did you choose?**

Region for *B. rapa* gene is Bra000234, chr A03: 10,034,947–10,137,601 and *B. oleracea* gene region Bol020542, chr C03: 12,069,323–12,172,042.

1. **Is there strong microsynteny? How much of the sequnence or gene models in that region are conserved?**

Microsynteny strength is really strong. There are large clearly conserved collinear blocks with some intergeneric spaces.

1. **Are there noticeable gaps, expansions, or rearangements between these two species? If so, what might be causing them?**

I can see gaps in intergenic regions. They could be due to gene loss during duplication, which also result in dotplot diagonal interruption. It may also be the outcome of transposable-element activity, inversions and transpositions.

1. **What is the sequence homology (e.g., average nucleotide conservation between the two species?**

From the reports I would say it is orthologs homology with percent identity of 94.84% (of the biggest chunk).