Construction of the ancestral angiosperm karyotype (AAK)

We used 15 current genomes of the highly diverged lineages to extract protochromosomes of AAK. We compared any two genomes and identified the highly syntenic dotplot that indicates shared chromosomes or chromosome-like 'synteny blocks' through chromosomal fusions. Here we only used a few examples of two-genome comparison to show how we identified the protochromosomes. It needs to be noted that each protochromosome has been identified across three current genomes during these two-genome comparisons.

(1) The homologous dotplot between *Am. trichopoda* (Atr) and *Illicium verum* (Ive) indicate that eight chromosomes (Atr10, Atr11, Atr4, Atr5, Atr6, Atr7, Atr8 and Atr9, surrounded by differently colored boxes) from the former species were shared by the latter species (Fig. 3), indicating that they might represent eight protochromosomes for both species.

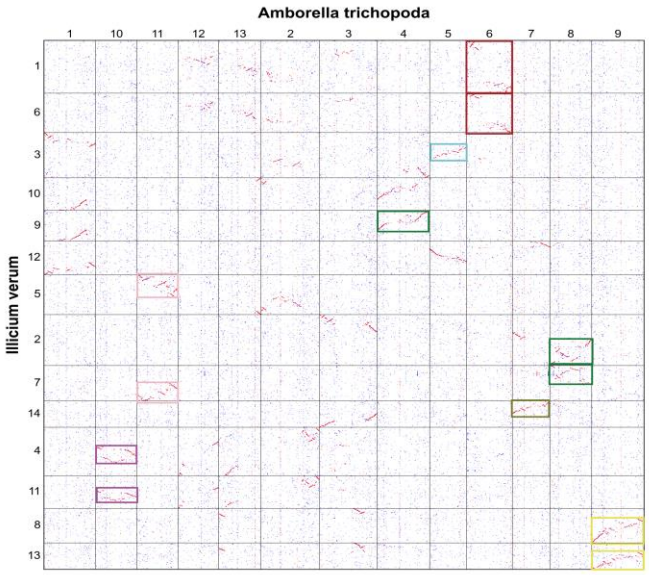

Figure 3 Genomic comparison between *Am. trichopoda* and *I. vernum* and the dotplot surrounded by differently colored boxes indicate the shared chromosomes or chromosome-like 'synteny blocks' through chromosomal fusions.

(2) The dotplot between *Am. trichopoda* and *L. chinense* (Lch) (Fig.4) show that Atr5, Atr6 and Atr8 are intact chromosomes, confirming the findings through a comparison between

Am. trichopoda and *I. verum* (Fig. 3). Furthermore, Lch19 is a proper subset of Atr2 (Fig.4), indicating it as a protochromosome for AAK.

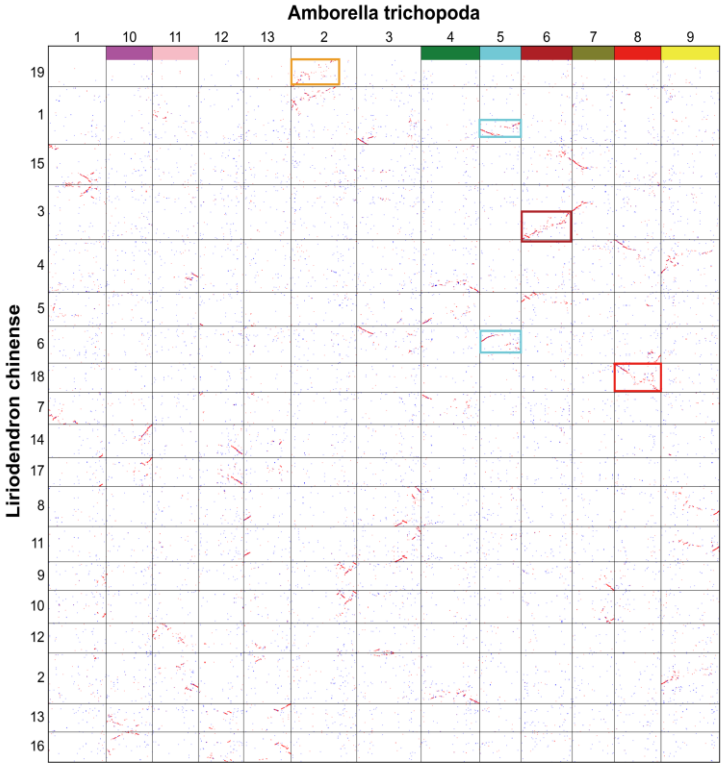

Figure 4 Genomic comparison between *Am. trichopoda* and *L. chinense* and the dotplot surrounded by differently colored boxes indicate the shared chromosomes or chromosome-like 'synteny blocks' through chromosomal fusions.

All of these nine protochromosomes are confirmed at least once in other two-genome comparisons. The next cycle of protochromosome identification can further use these two-genome comparisons but delete (or cover) the already identified protochromosomes.

(3) We can color (cover or delete) the already identified protochromosomes and revisit the dotplot between *Am. trichopoda* and *I. verum*. We use WGDI with the "-km,-d" parameters to complete this. According to the EEJ and NCF chromosomal fusions, if one protochromosome is the proper subset of a current chromosome, then the rest of this chromosome is also a protochromosome. The rest of Atr2 is also nearly intact in Ive4 and Ive11 (Fig. 5), which is likely to be one protochromosome for AAK. Similarly, after removing the already identified Atr4 as a protochromosome, the rest of Ive9 remains to be a nearly intact chromosome in Atr1, which should represent another protochromosomes for AAK. After removing Atr7 as one really identified protochromosome, the rest of Ive14

remains to be one nearly intact chromosome in Atr3 (Fig. 5), which likely represents another protochromosome for AAK.

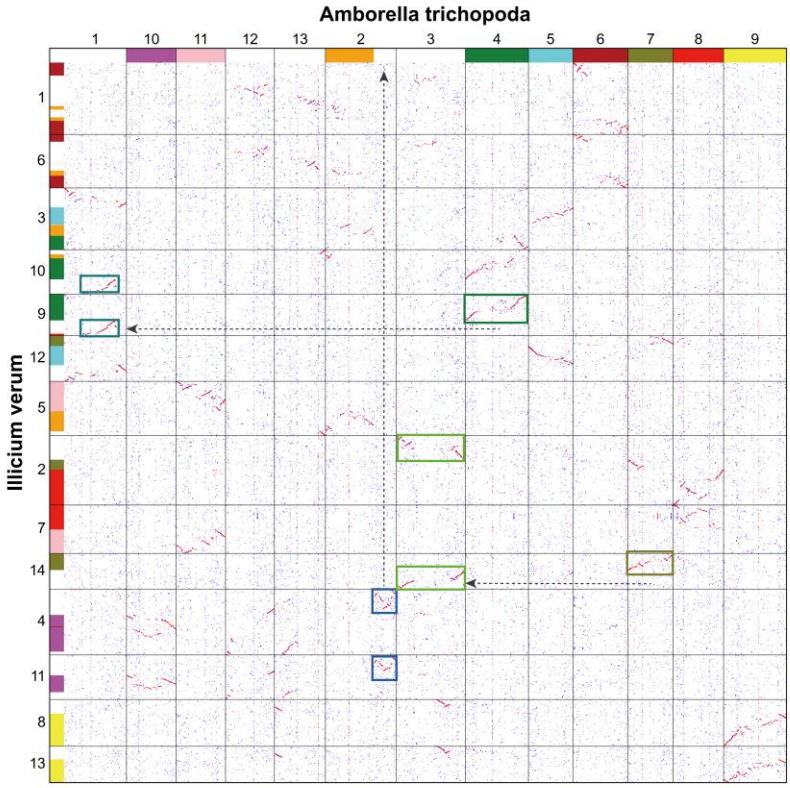

Figure 5 Further genomic comparison between *Am. trichopoda* and *I. verum* and the dotplot surrounded by differently colored boxes indicate the shared chromosomes or chromosome-like 'synteny blocks 'through chromosomal fusions after covering (or deleting) the already identified protochromosomes.

Further comparisons indicate that after deleting one already identified protochromosome, the rest of Atr1 is nearly intact in Ive3 and Ive12 (Fig. 6). This part is likely to represent another protochromosomes for AAK.

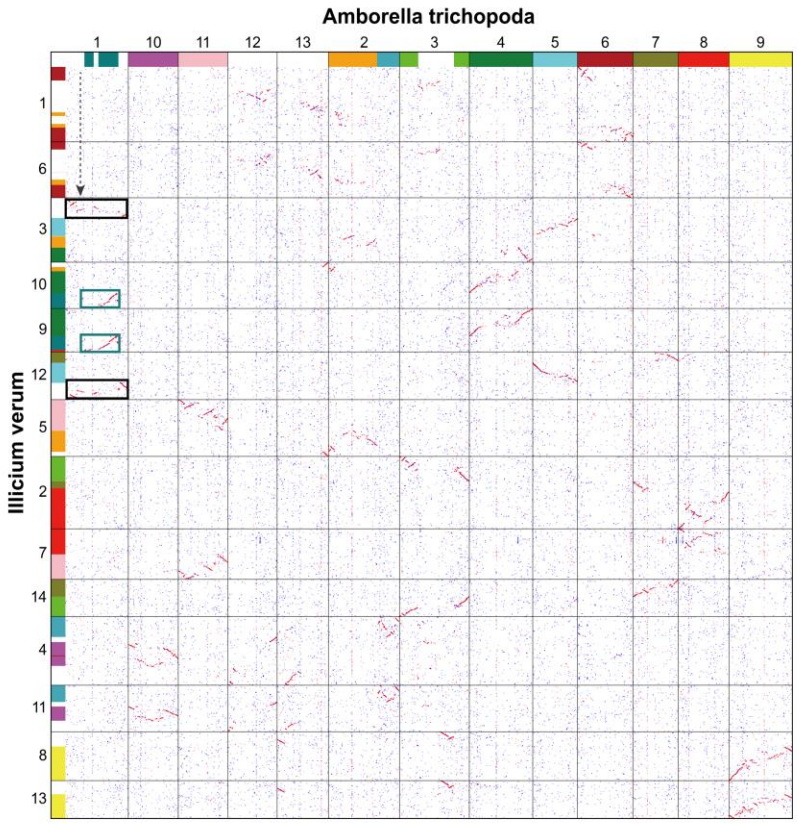

Figure 6 Further genomic comparison between *Am. trichopoda* and *I. verum* and after deleting one already identified protochromosome, the rest of Atr1 is nearly intact in Ive3 and Ive12. The part of Atr1 likely represents one protochromosome.

(4) After identifying and deleting the above-mentioned protochromosomes, we can find the rest of the chromosomes in *Am. trichopoda* and *I. verum*, Atr12, Atr13 and Atr3, (Ive8, Ive13) and (Ive4, Ive11) may represent other protochromosomes because the deleted parts are intact protochromosomes not fragments of protochromosomes.

The dotplot between *Am. trichopoda* (Atr) and *A. gramineus* (Agr) (Fig. 7) shows that Atr1(two protochromosomes), Atr4, Atr5, Atr6, Atr7, Atr8, Atr9, Atr10, Atr10 and Atr11 are intact and shared protochromosomes. The dotplot between *I. verum* (Ive) and *A. gramineus* (Agr) (Fig. 8) shows that (Agr5, Agr7) and (Ive8, Ive13) are intact and shared protochromosomes. After removing Atr9 as one identified protochromosome, the rest of Agr5 represents one protochromosome for AAK. After removing two protochromosome (Atr11 and the middle area of Atr1) (Fig. 7), the rest of Agr6 as chromosome-like 'synteny blocks' is nearly intact in Ive1 and Ive6 (Fig. 8), which are likely to represent one protochromosome for AAK.

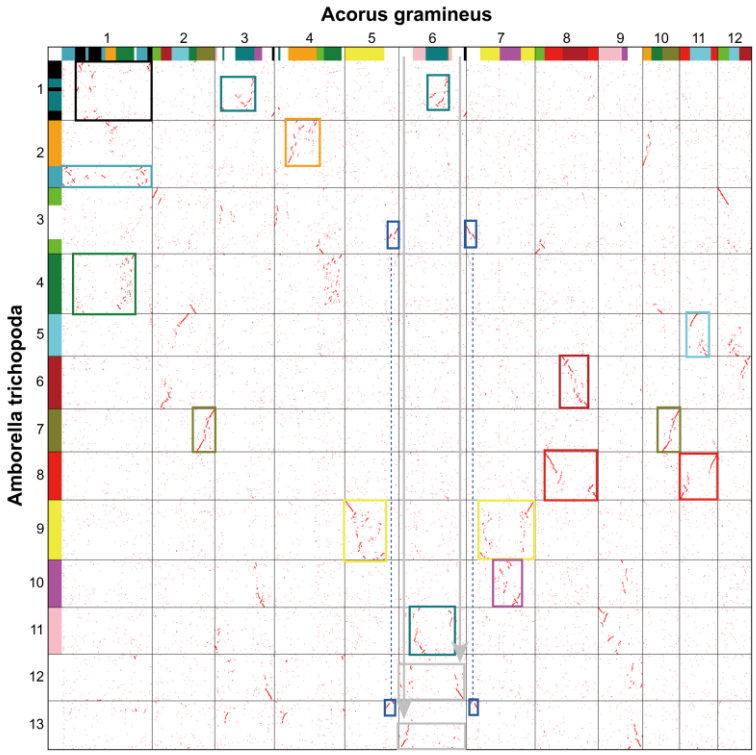

Figure 7 Genomic comparison between *Am. trichopod*a and *A. gramineus* and the dotplot surrounded by differently colored boxes indicate the shared chromosomes or chromosome-like 'synteny blocks' through chromosomal fusions.

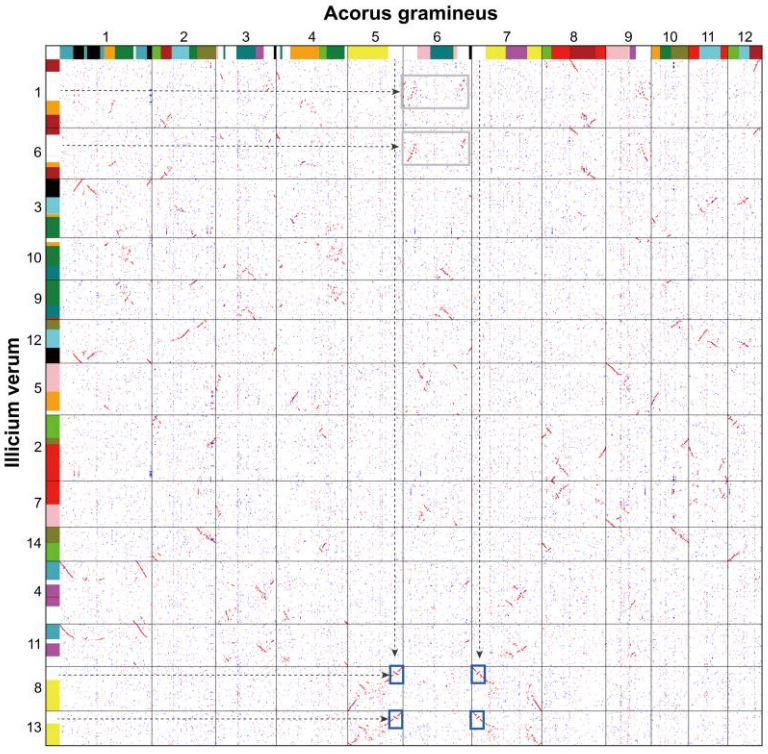

Figure 8 Genomic comparison between *I. verum* and *L. chinense* and the dotplot surrounded by differently colored boxes indicate the shared chromosomes or chromosome-like ‘synteny blocks’ through chromosomal fusions.

(5) The dotplot between *Am. trichopoda* (Atr) and *Nymphaea colorata* (Nco) (Fig. 9) shows that blank areas of *Am. trichopoda* are chromosome-like synteny blocks, which are likely to represent one protochromosomes for AAK. Furthermore, 14 of the 16 protochromosomes are chromosomes or chromosome-like synteny blocks in *N. colorata*.

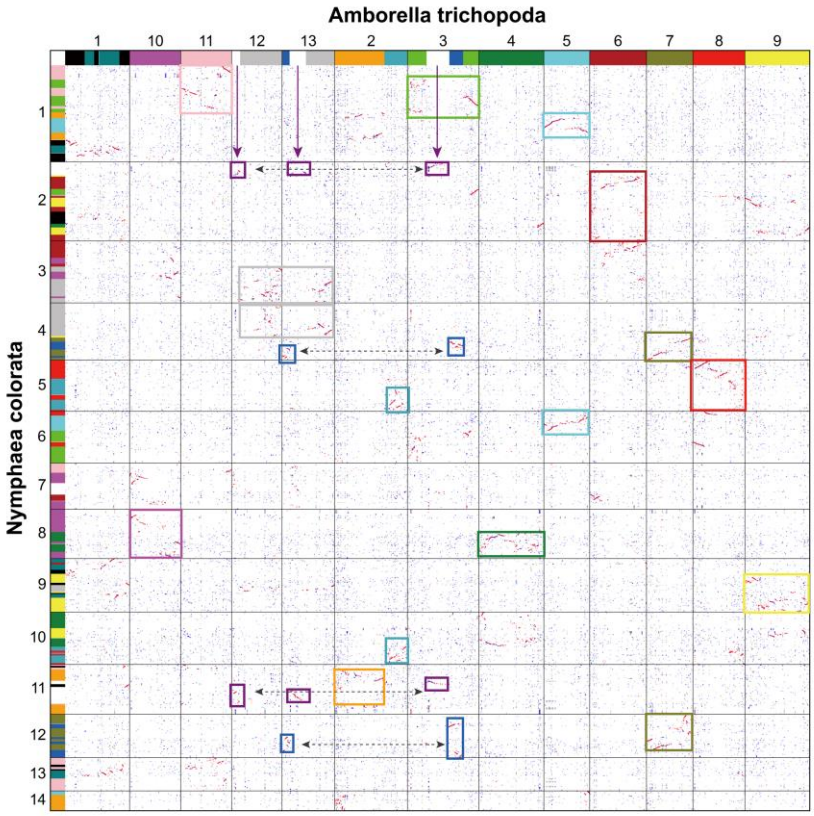

Figure 8 Genomic comparison between *N. colorata* and *Am. trichopod*a and the dotplot surrounded by differently colored boxes indicate the shared chromosomes or chromosome-like ‘synteny blocks’ through chromosomal fusions.

All of the remaining 7 protochromosomes in the second stage, can be identified at least once in other two-genome comparisons. We can use WGDI with the "-ak" parameter to obtain all protochromosomes for AAK. AAK comprises 16 protochromosomes according to our two-genome comparison of 15 current genomes from highly diverged lineages.

(6) We can further compare AAK and each current genome to determine the protochromosomal changes to evolve into the current chromosomes of one species. For

example, the dotplot between *Am. trichopoda* and AAK, we can identify three distinct chromosomal fusion events, NCF, RTA and NCF (Fig. 9).

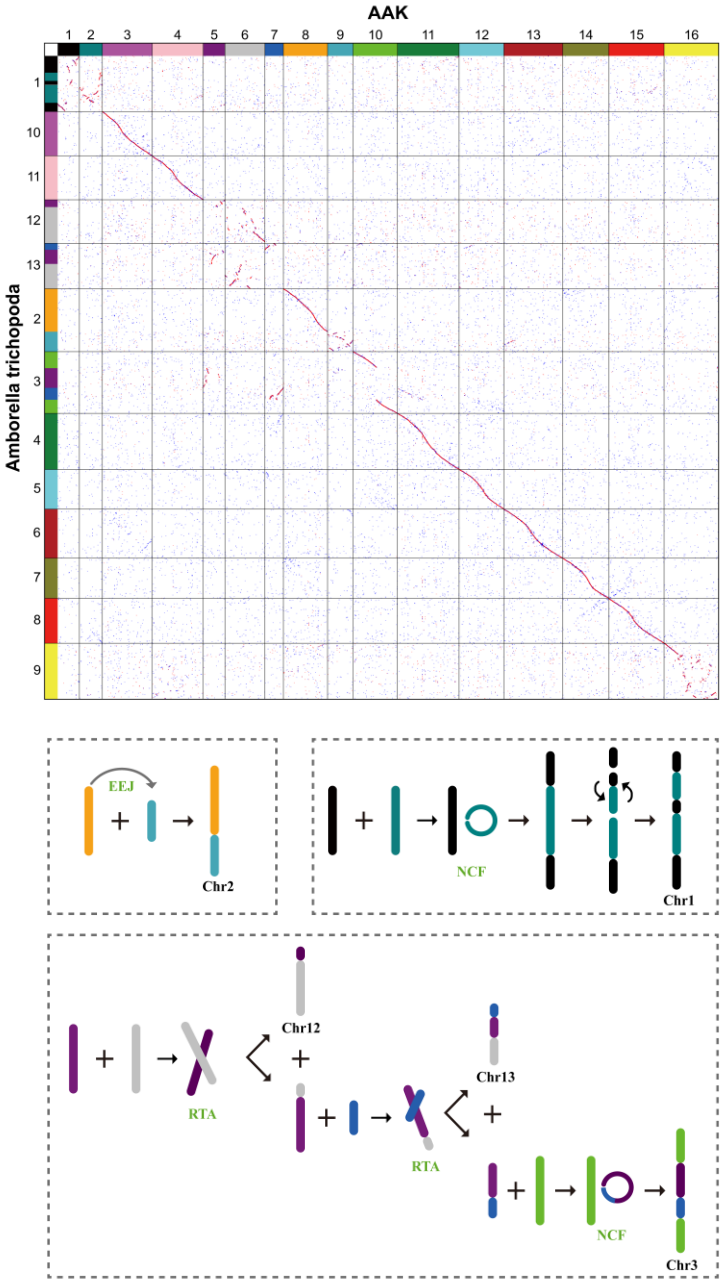

Figure 9. Genomic comparison between AAK and *Am. trichopoda* based on dotplot and the inferred chromosomal fusion events.