

Zhang-Lab 生信小课堂 第二十期

Applied Bioinformatics Club (ABC)

和趣求真  秉实生信

(张建伟生物信息学课题组 <https://zhang.hzau.edu.cn>)

Genome Puzzle Master(GPM)的使用

2025.03.07 二综C102 15:00 欢迎大家交流学习!

主讲人: 赵闻

2025/03/07



01 应用背景

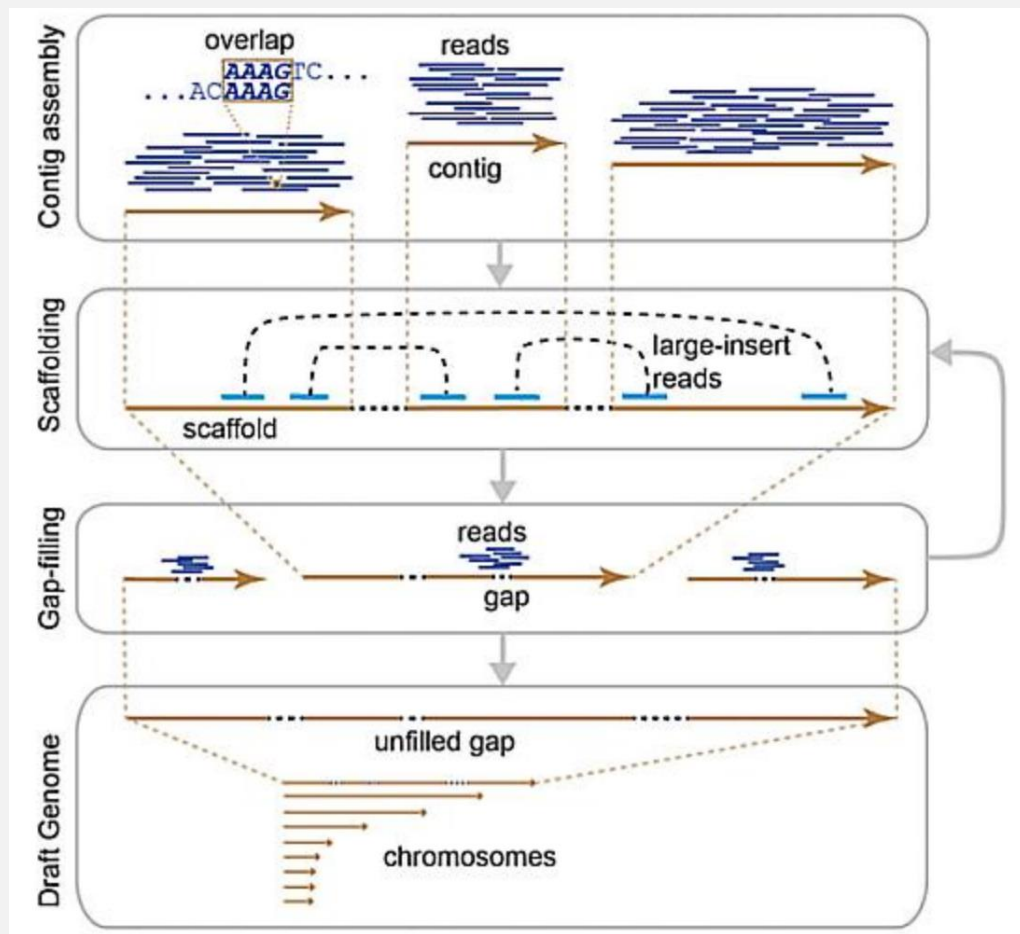
02 组装流程

03 操作展示



目录

1 应用背景



(Sohn and Nam 2018)

基因组组装目前主要分为两种方法：

- 映射比对组装 (Mapping assembly)
- 从头组装 (Denovo assembly)。

从头组装可以分为三步：

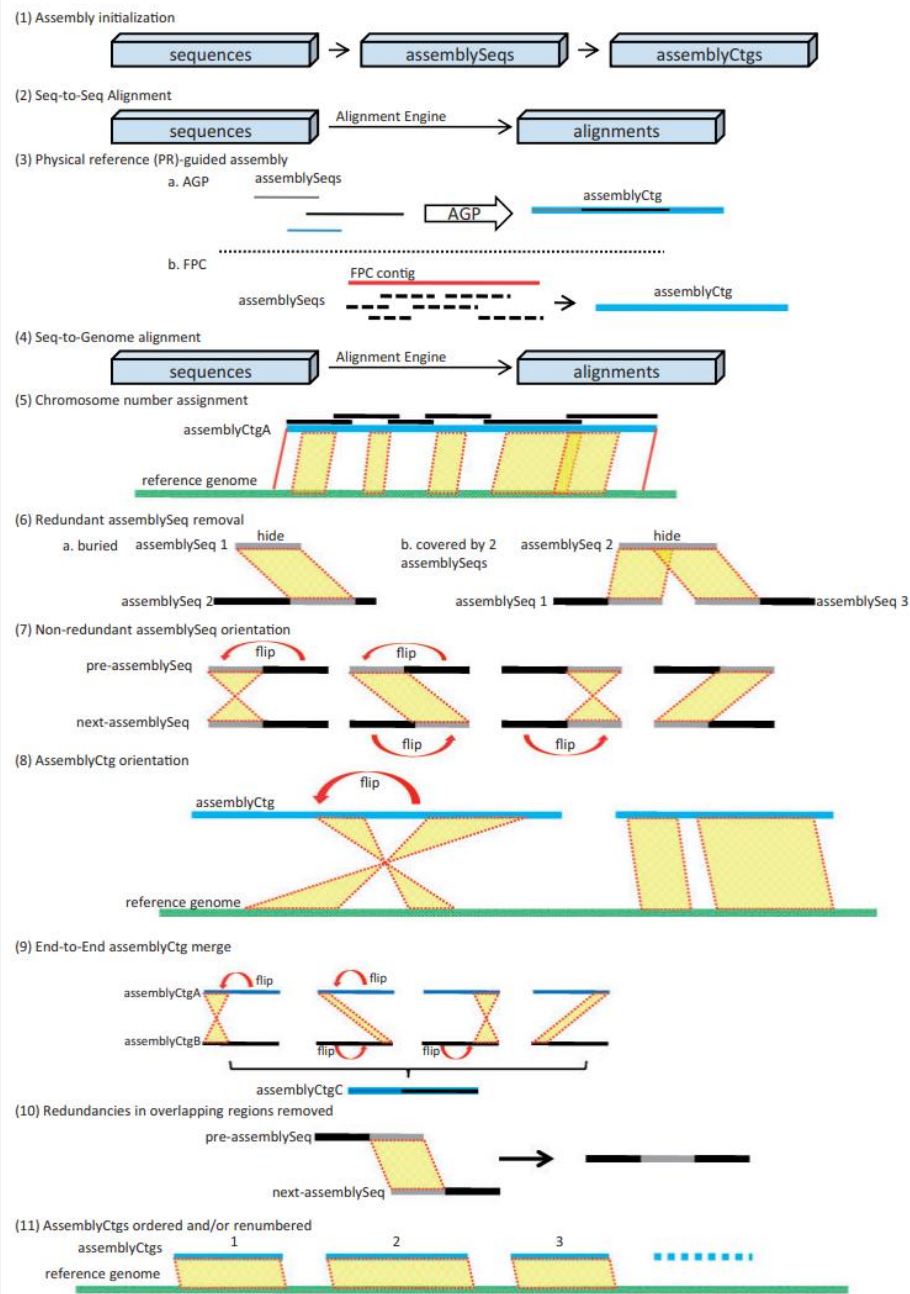
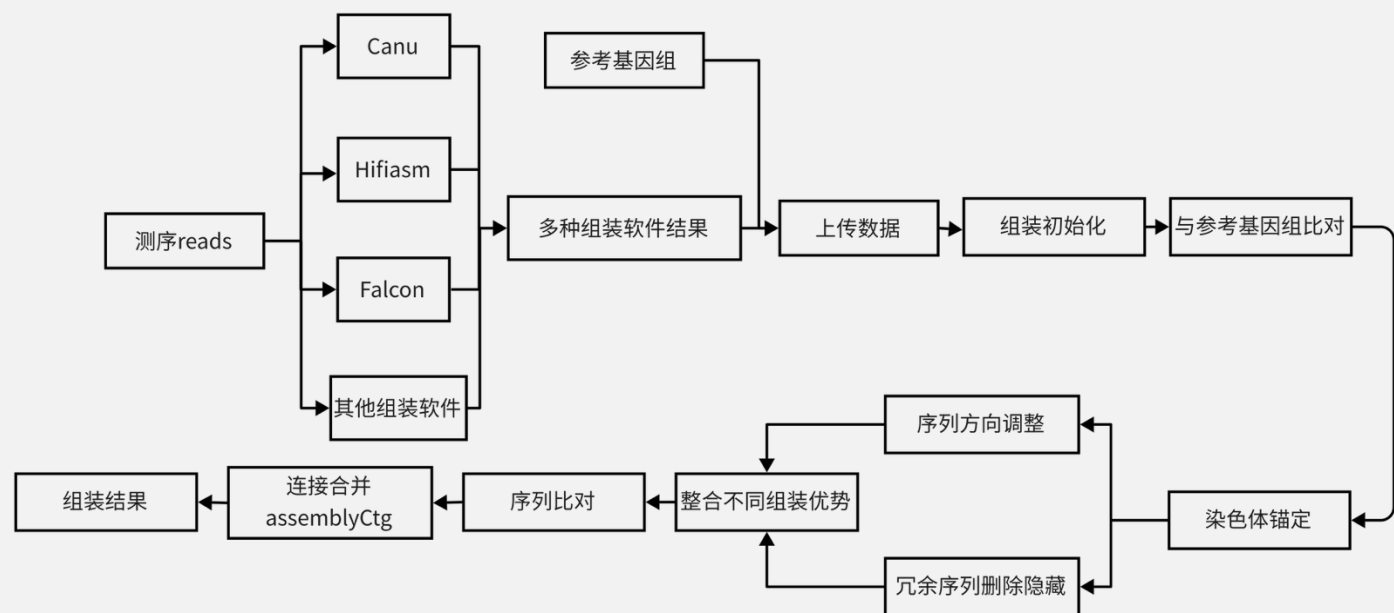
- (1) 测序获得的读长根据重叠关系，通过不同的拼接策略拼接成较长的重叠群；
- (2) 根据额外的信息（长读长、参考基因组信息等）判断重叠群的方向等相关信息，形成比重叠群更长的支架；
- (3) 通过高通量染色体构象捕获 (Hi-C) 信息、参考基因组以及光学图谱等辅助手段将支架挂载到染色体上，生成染色体水平的基因组。

基因组组装工具：RagTag、QuickMerge、GPM...

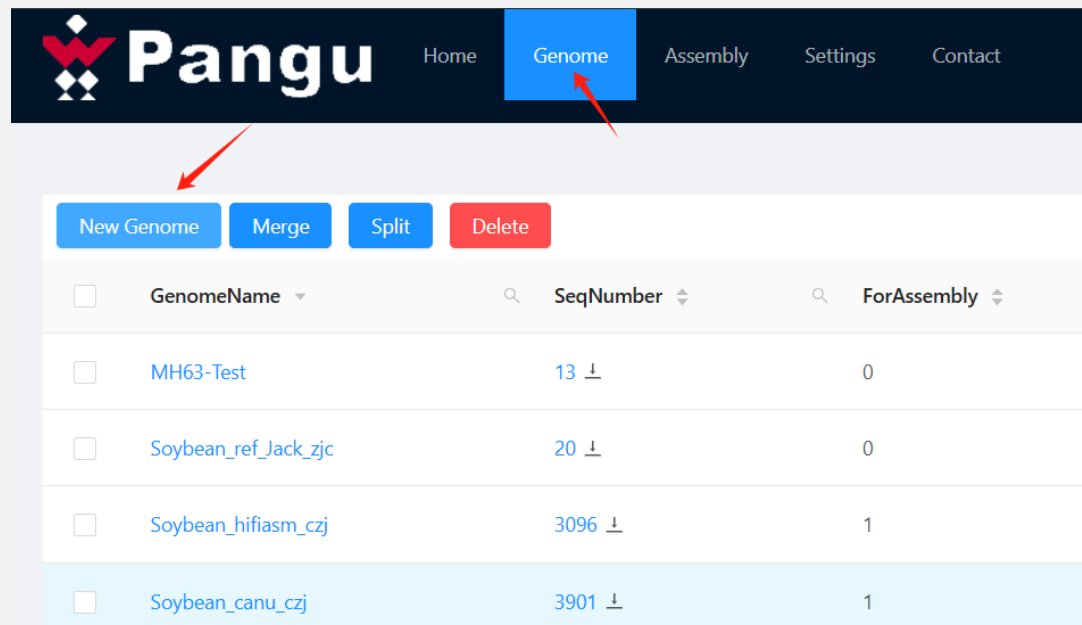
GPM网址：

<https://riceome.hzau.edu.cn/gpm/#/login>

2 组装流程

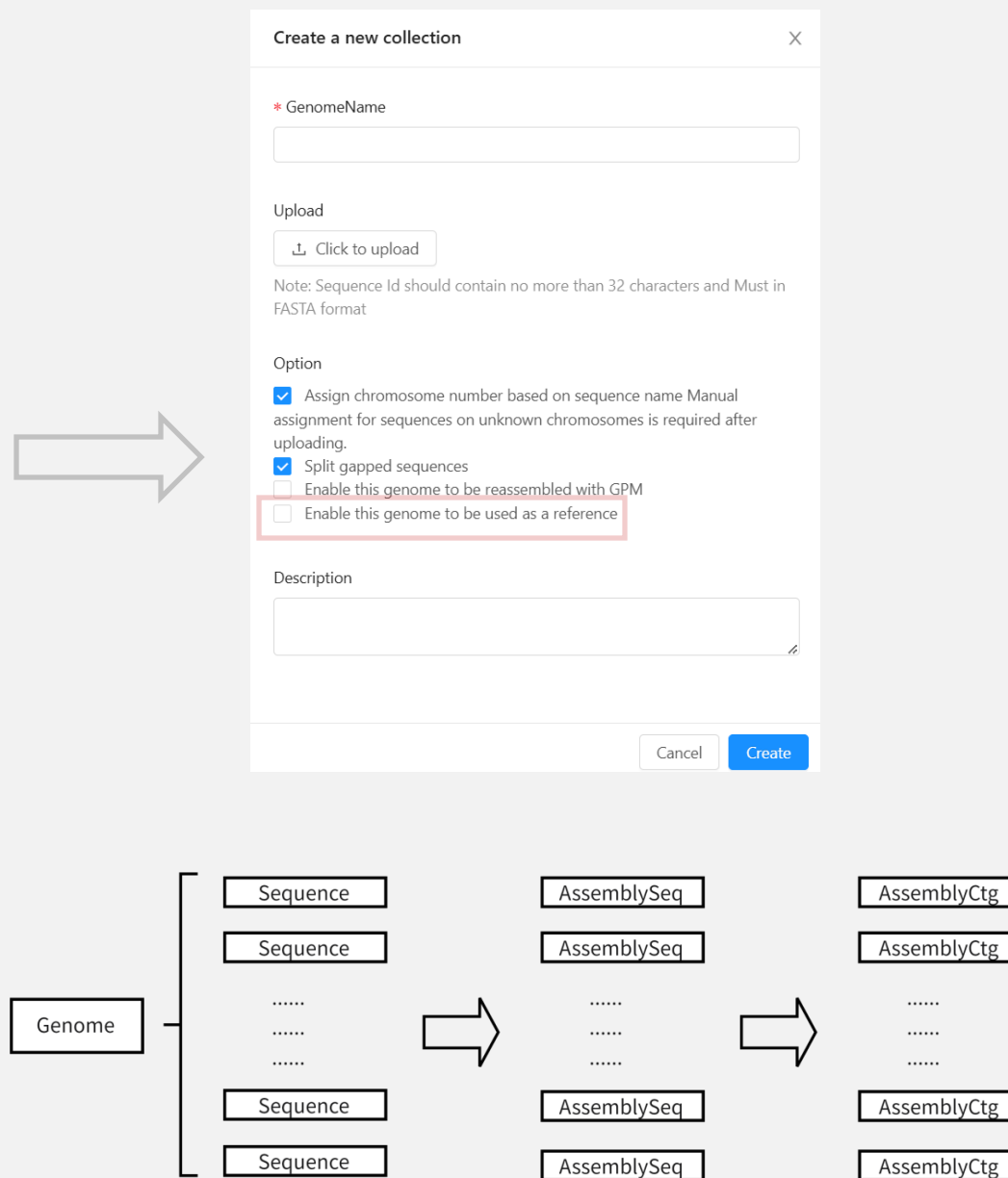


3 操作展示-组装初始化



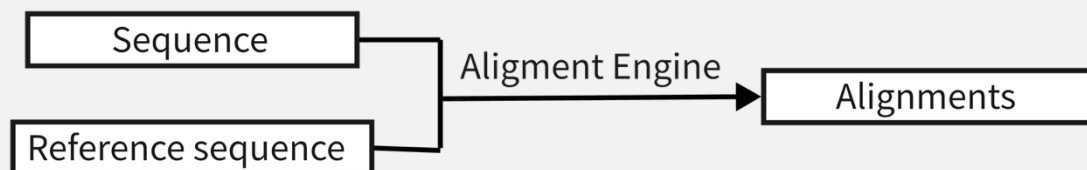
点击Genome—New Genome上传参考基因组与组装软件的.fasta文件。

将组装软件初步组装获取的contig初始化为assemblySeqs，并以每个assemblySeq作为源序列初始化一个assemblyCtg，形成一个组装的独立数据。

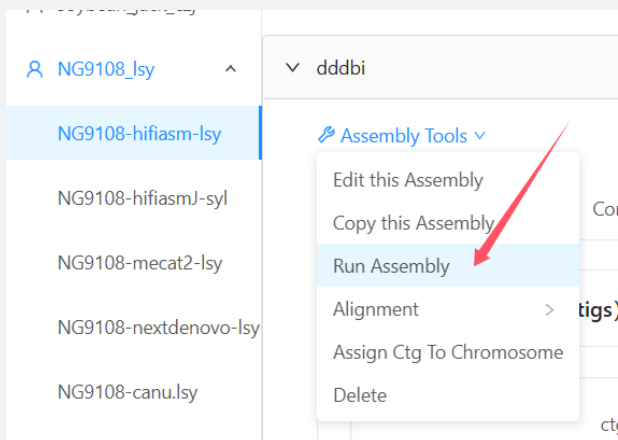


3 操作展示-比对到参考基因组

将参考基因组与组装软件的组装结果比对，获取比对信息



3 操作展示-染色体锚定



Run Assembly

Start Over a New Assembly?: ☐ Yes ☒ No

With All Sequences available?: ☒ Yes ☐ No

AssemblySeq Minimum Length:

Referenc Genome:

Alignment Block Size(bp):

Alignment Coverage Percent(%):

Alignment Settings:

- ☒ Assign chromosome number for unplaced contigs
- ☐ Re-estimate position for anchored contigs
- ☒ Orient contigs based-on reference genome

End-to-End Merge: ☐ Yes ☒ No

Minimum Overlap:

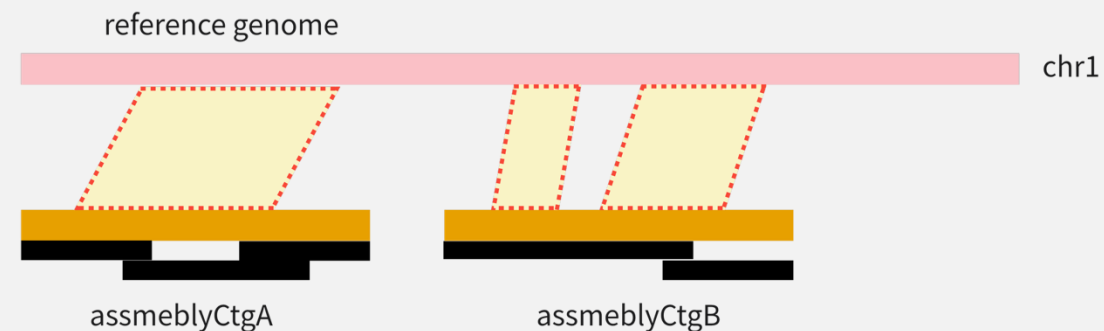
Minimum Identity:

Filter: ☒ Redundant Sequences ☒ Redundant Overlaps

Auto-Orient Sequences: ☒ Yes ☐ No

Renumber Contigs: ☒ Yes ☐ No

Cancel OK



NG9108_hifiasm_lsy

Assembly Tools

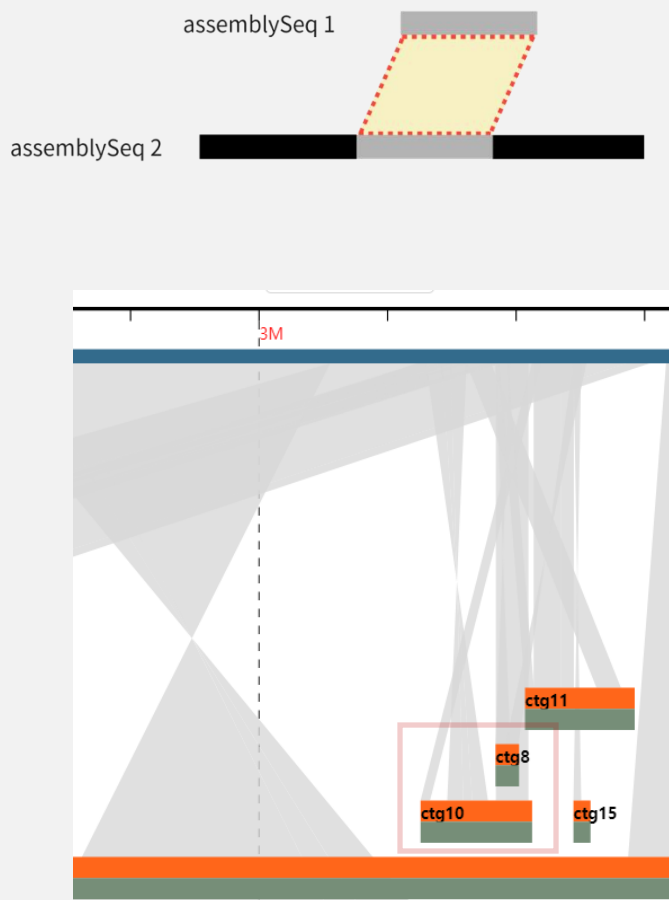
Assembly(1421) Contig List Stats Check New Sequences About

Chromosome 1 (2 Contigs) - 2 Seqs 44,086,576bp	
ctg1421 365,132 bp	ctg1 43,721,444 bp

Chromosome 2 (1 Contigs) - 1 Seqs 36,499,651bp	
ctg1227 36,499,651 bp	

Chromosome 3 (1 Contigs) - 1 Seqs 37,903,698bp	
ctg99 37,903,698 bp	

3 操作展示-冗余序列调整



ctg8与ctg10比对块重叠冗余

Assign Ctg to Chromosome

Check All Uncheck All

Please check ctgs for assigning chromosome number:

Start Range End Range Select Range

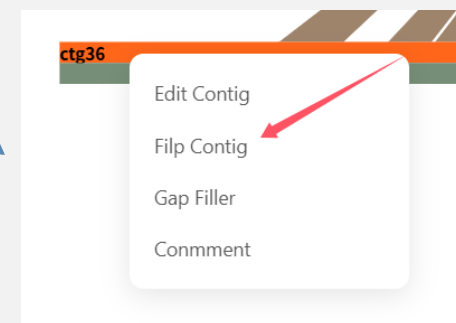
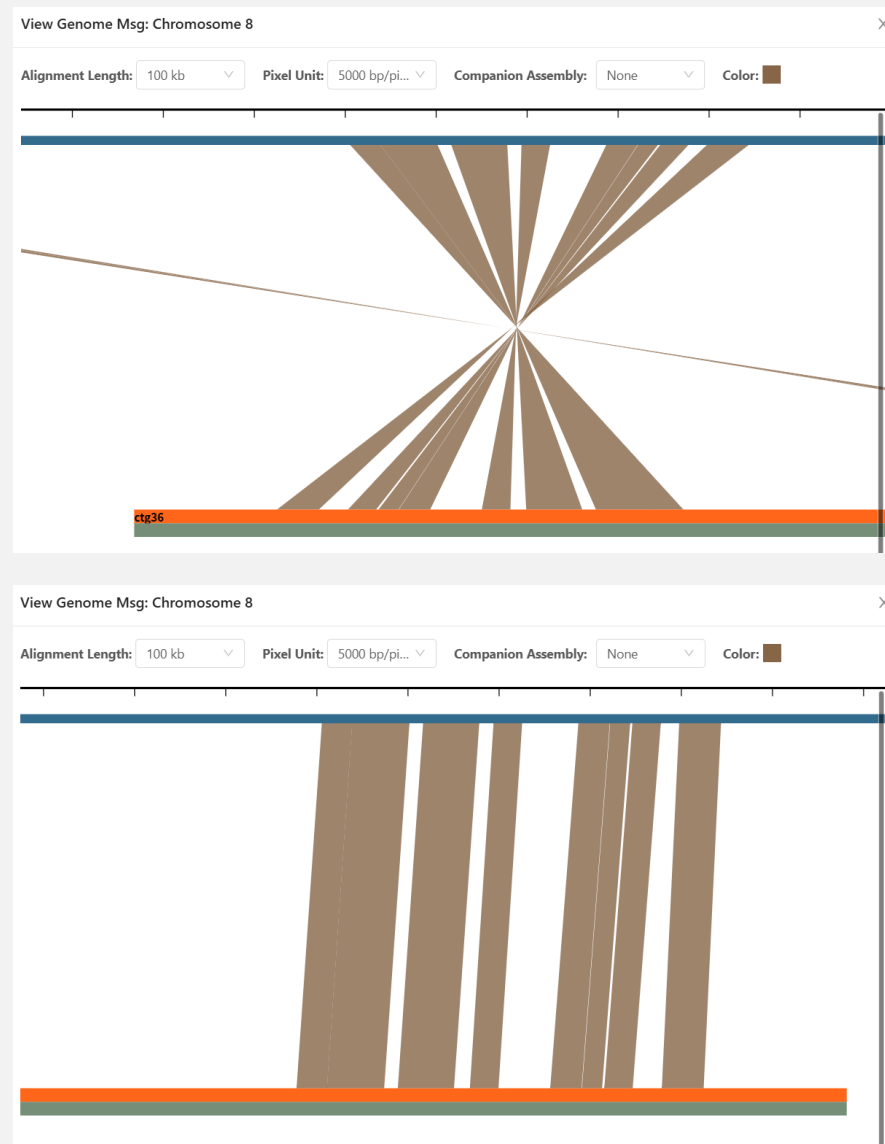
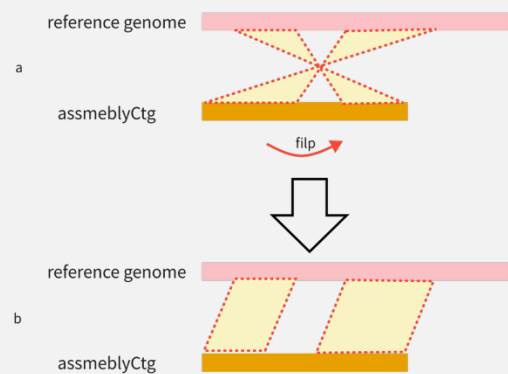
☐ ctg1	☐ ctg2	☐ ctg3	☐ ctg4	☐ ctg5	☐ ctg6
☐ ctg7	☒ ctg8	☐ ctg9	☐ ctg10	☐ ctg11	☐ ctg12
☐ ctg13	☐ ctg14	☐ ctg15	☐ ctg16	☐ ctg17	
☐ ctg18	☐ ctg19	☐ ctg20	☐ ctg21	☐ ctg22	
☐ ctg23	☐ ctg24	☐ ctg25	☐ ctg26	☐ ctg27	
☐ ctg28	☐ ctg29	☐ ctg30	☐ ctg31	☐ ctg32	
☐ ctg33	☐ ctg34	☐ ctg35	☐ ctg36	☐ ctg37	
☐ ctg38	☐ ctg39	☐ ctg40	☐ ctg41	☐ ctg42	
☐ ctg43	☐ ctg44	☐ ctg45	☐ ctg46	☐ ctg47	
☐ ctg48	☐ ctg49	☐ ctg50	☐ ctg51	☐ ctg52	
☐ ctg53	☐ ctg54	☐ ctg55	☐ ctg56	☐ ctg57	
☐ ctg58	☐ ctg59	☐ ctg60	☐ ctg61	☐ ctg62	
☐ ctg63	☐ ctg64	☐ ctg65	☐ ctg66	☐ ctg67	
☐ ctg68	☐ ctg69	☐ ctg70	☐ ctg71	☐ ctg72	
☐ ctg73	☐ ctg74	☐ ctg75	☐ ctg76	☐ ctg77	
☐ ctg78	☐ ctg79	☐ ctg80	☐ ctg81	☐ ctg82	
☐ ctg83	☐ ctg84	☐ ctg85	☐ ctg86	☐ ctg87	
☐ ctg88	☐ ctg89	☐ ctg90	☐ ctg91	☐ ctg92	
☐ ctg93	☐ ctg94				

Chromosome #: 0

Cancel OK

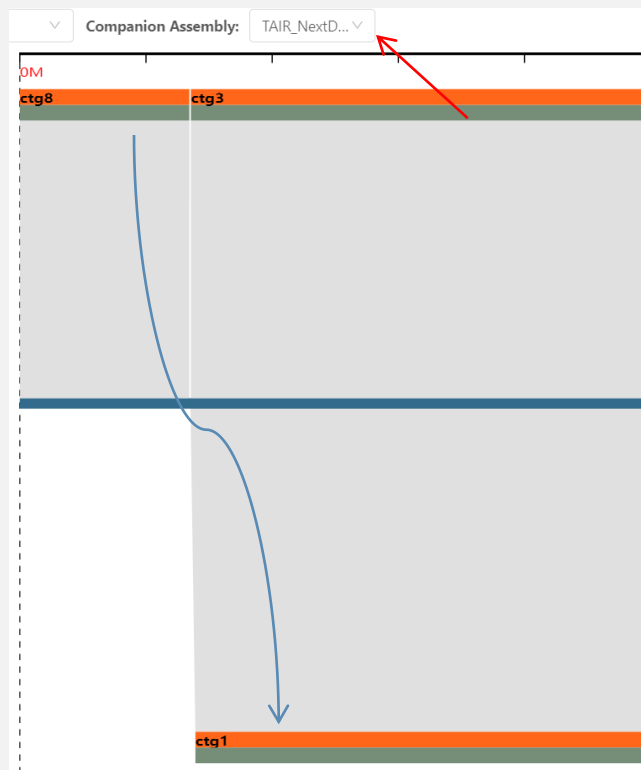
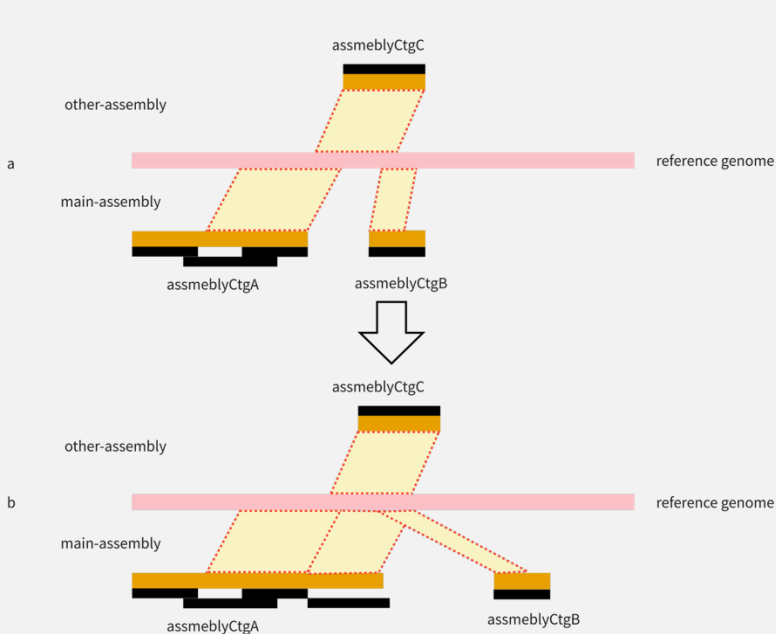
- 推荐用户将冗余的ctg移动到0号染色体上，不建议删除以防需要重新使用。
- 操作：Assembly Tools-Assgin Ctg to Chr

3 操作展示-序列方向调整



3 操作展示-不同组装优势整合

选择Companion Assembly即可比较不同软件组装结果



右击被添加的ctg8-View Details, 复制Assembly Id

View Sequence

Seq Name: contig_54

AssemblySeq Id: 97916

Assembly: < ctg10-NG9108 >

右击ctg1-Edit contig

GPM AssemblyCtg Editor:ctg7

Ctg Id: 98047

Chromosome: 12

Estimated Position: -1066666

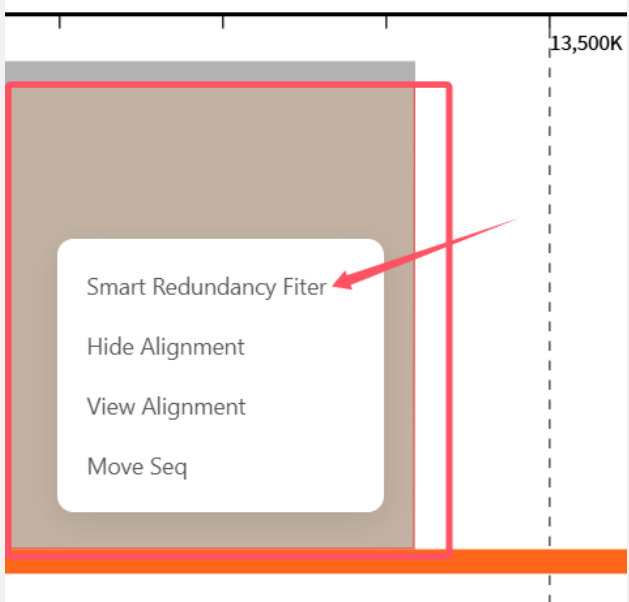
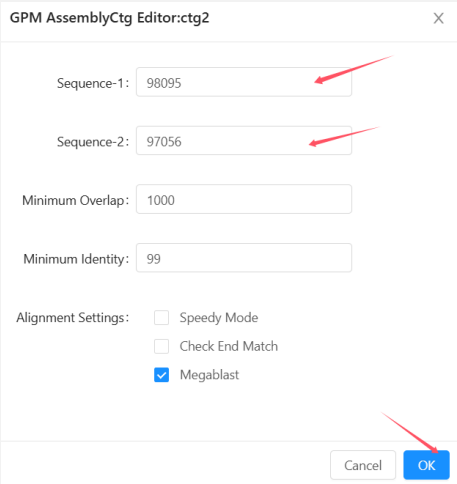
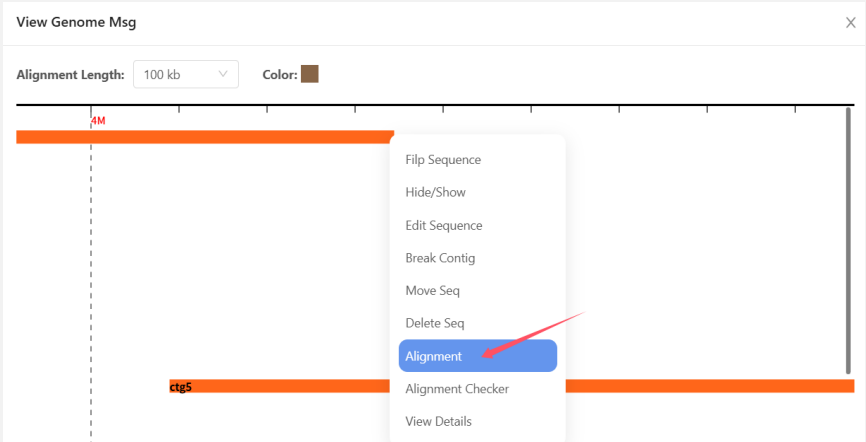
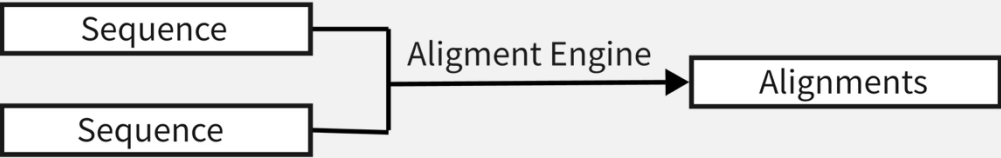
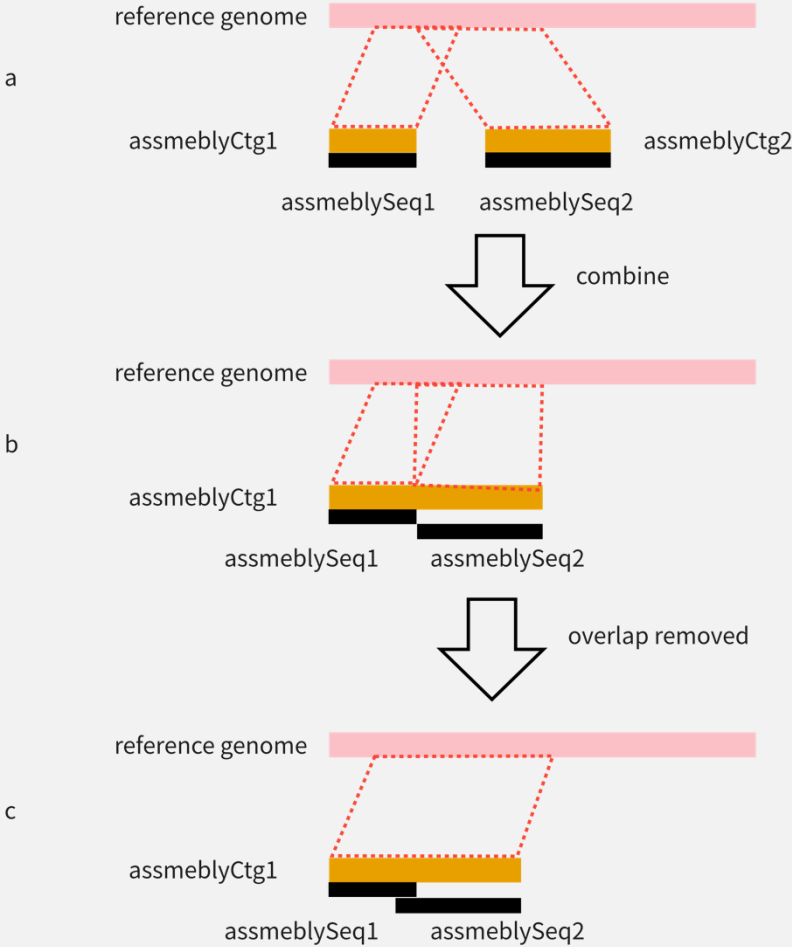
Sequences: 000112F

Insert Sequence: 97916

Options: ☐ Append to ctg7
☐ Flip the selected contig

Cancel OK

3 操作展示-序列比对



3 操作展示-组装结果

完成可视化组装后，可直接下载组装后的各染色体文件。

Isy-nextdenovo

Assembly Tools

Assembly(50)

Contig List

Stats

Check New Sequences

About

Chromosome 1 (1 Contigs) - 1 Seqs 43,448,084bp

ctg1

43,448,084 bp

Chromosome 2 (2 Contigs) - 2 Seqs 36,158,314bp

ctg6

1,855,814 bp

ctg7

34,302,500 bp

download

Ctg Pseudomolecules

Chr Pseudomolecules

3 操作展示

拖拽一个ctg下的seqs可以随意调换seq间的顺序。

GPM AssemblyCtg Editor:ctg70

Ctg Id: 98091

Chromosome:

4

Estimated Position:

21622399

Sequences:

ctg5

000156F

Insert Sequence:

Insert Seq

Options:

☐ Append

to ctg70

☐ Flip the selected contig

Cancel

OK

组装状态统计

Assembly Tools

Assembly(50)

Contig List

Stats

Check New Sequences

About

Assignment	Number of Contigs	Length (bp)	Component Seqs
Chromosome 1	1	43,448,084	1
Chromosome 2	2	36,158,314	2
Chromosome 3	2	37,602,934	2
Chromosome 4	1	14,542,112	1
Chromosome 5	2	30,099,219	2
Chromosome 6	6	12,392,977	6
Chromosome 7	1	29,680,954	1
Chromosome 8	2	22,001,615	2
Chromosome 9	4	22,889,243	4



感谢观看

汇报人：赵闻