

Phyphya เป็น lightweight python lib

↳ predicts ระบุความหมายใน analyzing dataset ก่อนเริ่มทำการ

ML-based tree inferences

↳ using a statistical method known as Maximum Likelihood (ML)  
↳ เพื่อสร้าง Phylogenetic tree

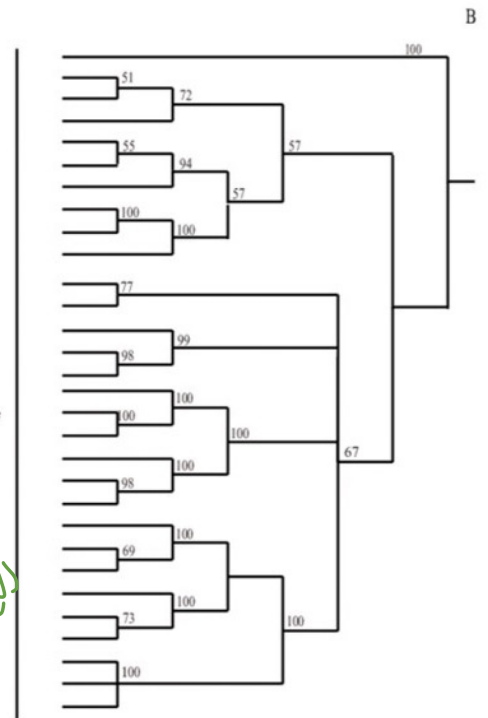
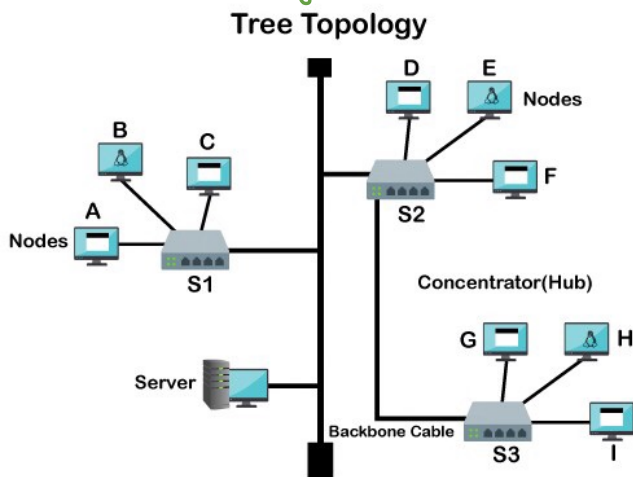
Phylogenetic trees

• ถูกนำมาใช้ใน biology (evolutionary relationships)  
ex. genes, proteins

• จุดมุ่งหมายคือการหา tree topology

↳ เครือข่ายมี trunk & node ตรงกลาง แต่ละโหนดเชื่อม  
ต่อกับโหนดกลางผ่านเส้นทางการเดินทาง node สามารถมอง  
เห็นว่ามีมาจาก branches จาก trunk ใช้ในเครือข่ายใหญ่

↳ branch lengths



- sequence alignments

↳ character data  
↳ parameters

- จัดการข้อมูล (ดึง, ล้าง, แปลงรูปแบบ, รวมฐานข้อมูล)

- จัดตำแหน่ง (sequence)

- Model Parameter (numerical optimization)

- Visualization (tree diagram)

- Model selection เพื่อเขียน software

ข้อดีของ Pythia ต่อ phylogenetic analyses

1. Speed of prediction เพราะ predict ก่อนวิเคราะห์ data มีประสิทธิภาพ

2. Resource Efficiency ลดจำนวนการเลือกใช้ทางเลือกในการวิเคราะห์

3. workflow ที่ถูกปรับให้เหมาะสม

4. Improved Decision-Making เพราะ Pythia แจ้งขั้นตอนการวิเคราะห์ phylogenetic  
เพื่อเป็นแนวทางในการ alignment methods, substitution models, การแบ่ง partition

Phyia สามารถเพิ่มการรับรู้ของผู้ใช้เกี่ยวกับปริมาณ signal และ uncertainty expect ใน phylogenetic จีโนมการแก้ (post-)analysis setup ที่เหมาะสม (เพื่อประเมินคุณภาพ dataset)

- ประเมินสัญญาณ ใน phylogenetic information เพื่อ check ความแตกต่างของคุณภาพ และข้อมูล
- choice of Search Algorithms เพื่อเลือก tree search algorithms ที่ดีที่สุด

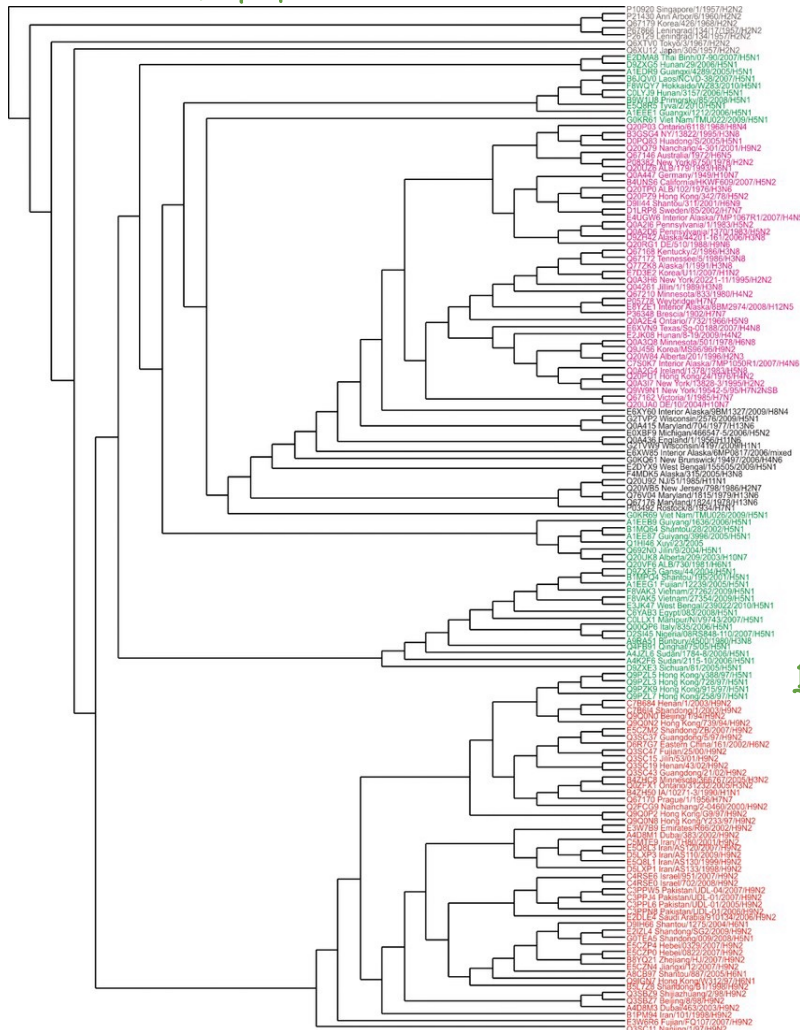
Easy-to-Analyze Datasets เลือก Maximum Likelihood (ML) ex. RAxML-NG

Intermediate-to-Analyze Datasets พิจารณาประสิทธิภาพและความแม่นยำ PhyML

Hard-to-Analyze Datasets สัญญาณต่ำ ความไม่แน่นอนสูง Phythia เน้นวิธีการ IQ-TREE และ จัด Algorithms หรือปรับปรุงข้อมูลก่อนการวิเคราะห์

Phyia supports ข้อมูลในรูปแบบ

phylip



FASTA format

```
>AT1G09780|1|training
GTGGAGTAGAAGAATTGAGAGCCTTATCAG
TTTTTGAAGAGAGGGCTGAACTCTCTAGT
TATCTTTTGTGCTTTTCTAATAATAAGAG
TTTACACACAG
>AT1G31812|0|testing
TCCTCATCTGCAGTAACCTTTATCTTAAGCA
TCAAATAACATTGCATAAGACTTGTTCCTT
GCTCTTGTGTTTCTATCATATTTAAGCTAT
CTACTTTGTGA
```

Part 1

Part 2

Part 3

ใช้จัดเก็บ, แบ่งปันข้อมูล biological sequence data

1. tree construction

2. tree Visualization

3. Boot strap Analysis ใช้คำนวณ statistic เพื่อดู uncertainty tree topology

4. ทรจํานน tree

5. Data format Conversion

PHYLogeny Inforence Package  
-ใช้สร้าง phylogenetic trees