

# Output results of CLIME (CLustering by Inferred Models of Evolution)

## Dataset:

Num of genes in input gene set: 13  
Total number of genes: 20834  
Prediction LLR threshold: 0

The CLIME PDF output two sections:

### 1) Overview of Evolutionarily Conserved Modules (ECMs)

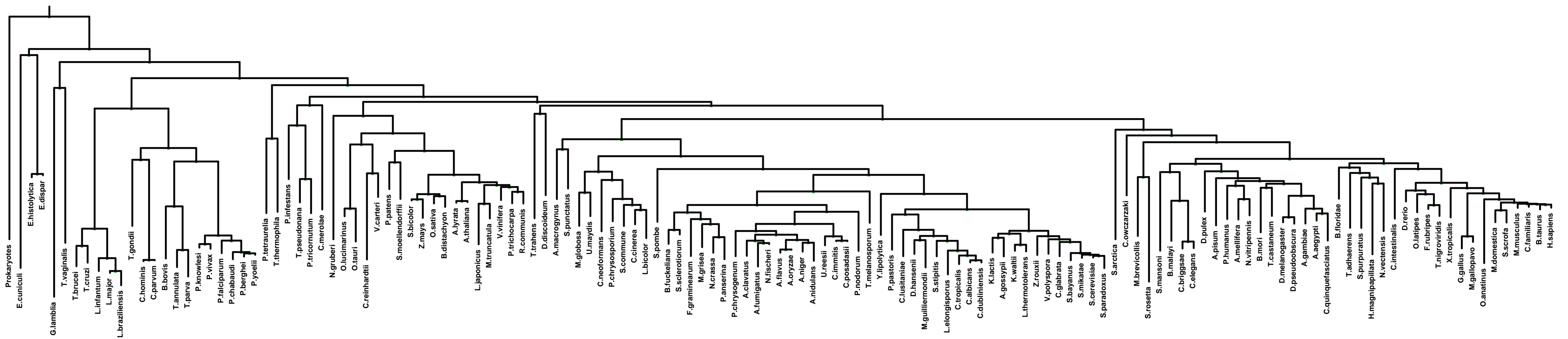
- Top panel shows the predefined species tree.
- Bottom panel shows the partition of input genes into Evolutionary Conserved Modules (ECMs), ordered by ECM strength (shown at right), and separated by horizontal lines.
- Each row show one gene, where the phylogenetic profile indicates presence (blue) or absence (gray) of homologs in each species (column).
- Gene symbols are shown at left. Gray color indicates that the gene is a paralog to a higher scoring gene within the same ECM (based on BLASTP  $E < 1e-3$ ).

### 2) Details of each ECM and its expansion ECM+

- Top panel shows the inferred evolutionary history on the predefined species tree. Branch color shows the gain event (blue) and loss events (red color, with brighter color indicating higher confidence in loss). Branches before the gain or after a loss are shown in gray.
- Bottom panel shows the input genes that are within the ECM (blue/white rows) as well as all genes in the expanded ECM+ (green/gray rows). The ECM+ includes genes likely to have arisen under the inferred model of evolution relative to a background model, and scored using a log likelihood ratio (LLR).
- PG indicates "paralog group" and are labeled alphabetically (i.e., A, B). The first gene within each paralog group is shown in black color. All other genes sharing sequence similarity (BLAST  $E < 1e-3$ ) are assigned to the same PG label and displayed in gray.

# Overview of Evolutionarily Conserved Modules (ECMs)

Last Common Ancestor



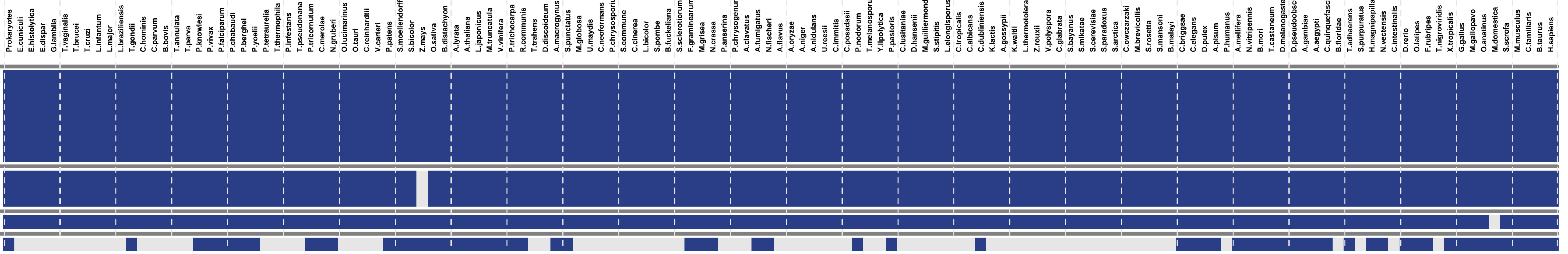
Protists

Plants

Fungi

Metazoa

Protein



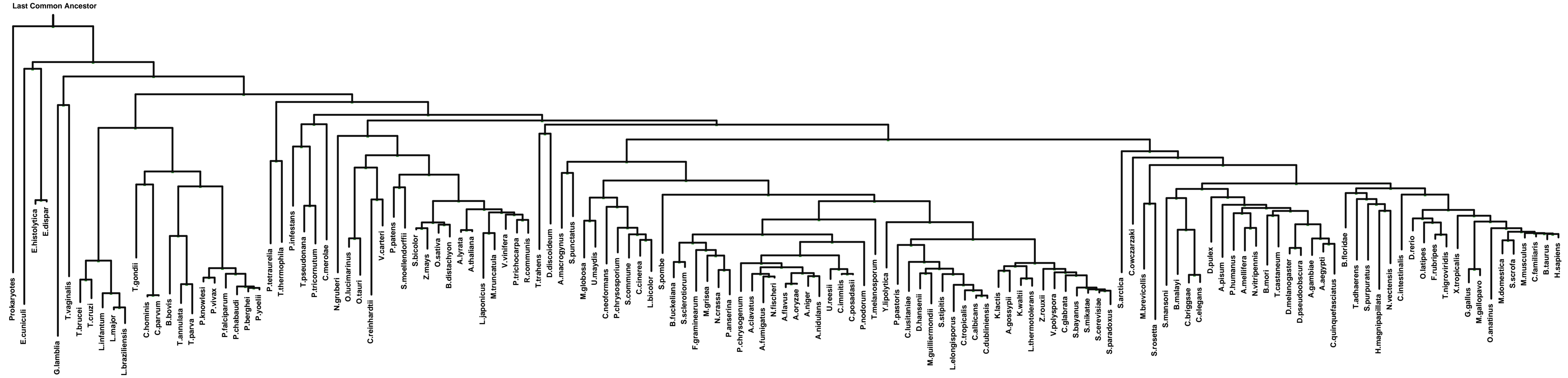
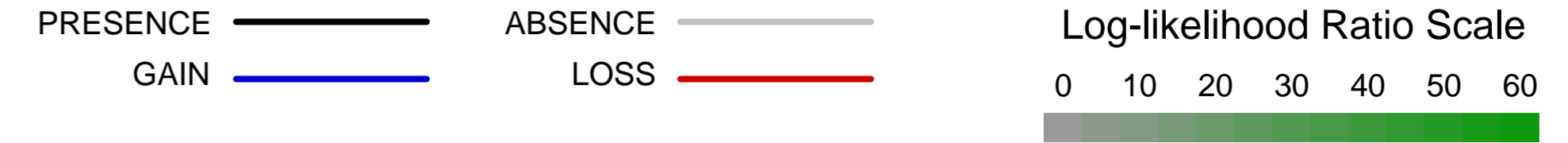
Strength

0.0

0.0

ECM 1, Gene set "Mismatch repair", Page 1

Num of ECM Genes: 8. Num of Predicted Genes: 0. ECM Strength: 0.0

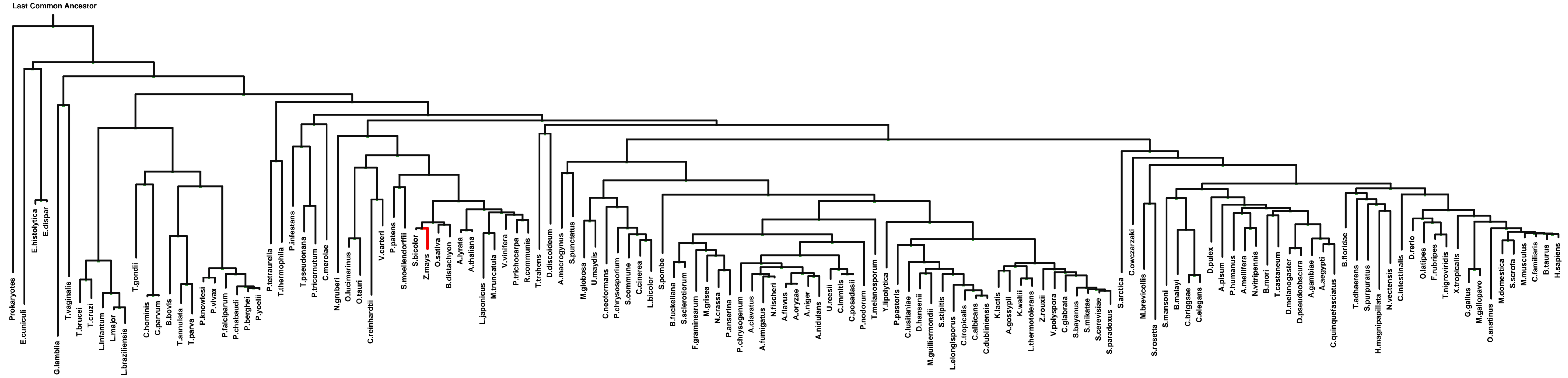


| PG | Protein | Prokaryotes | Protists | Plants | Fungi | Metazoa | LLR | Notes |
|----|---------|-------------|----------|--------|-------|---------|-----|-------|
| A  | EXO1    |             |          |        |       |         |     | 1     |
| B  | MSH2    |             |          |        |       |         |     | 1     |
| B  | MSH3    |             |          |        |       |         |     | 1     |
| B  | RFC1    |             |          |        |       |         |     | 1/2/3 |
| B  | RFC2    |             |          |        |       |         |     | 1/2/3 |
| B  | RFC3    |             |          |        |       |         |     | 1/2/3 |
| B  | RFC4    |             |          |        |       |         |     | 1/2/3 |
| B  | RFC5    |             |          |        |       |         |     | 1/2/3 |

1: Mismatch repair || 2: Nucleotide excision repair || 3: DNA replication

ECM 2, Gene set "Mismatch repair", Page 1

Num of ECM Genes: 3. Num of Predicted Genes: 0. ECM Strength: 0.0

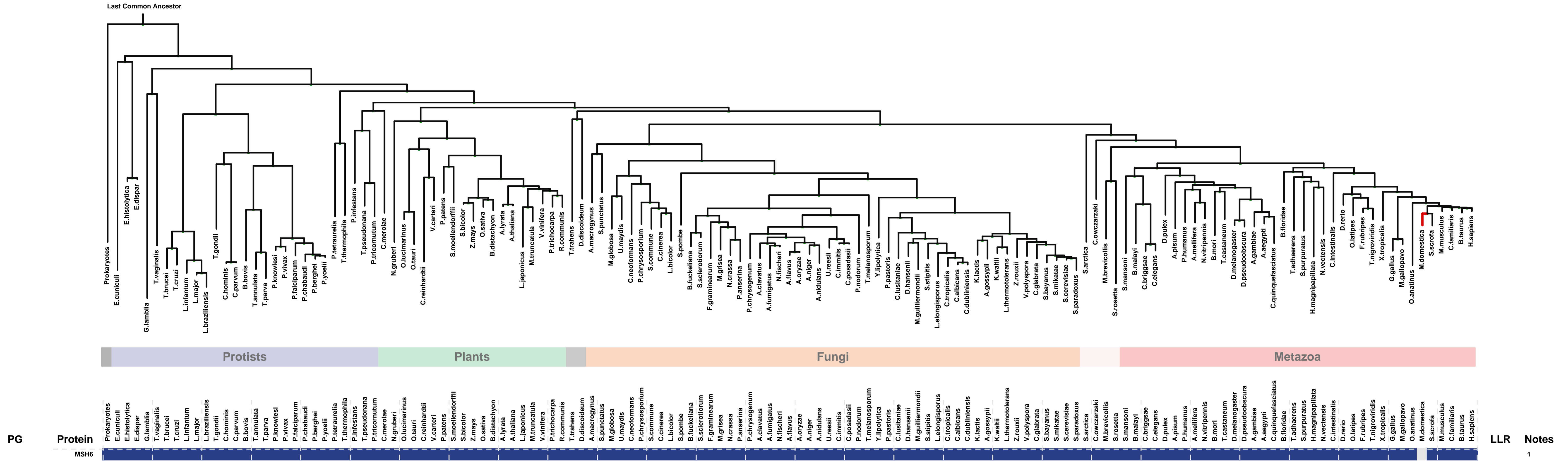
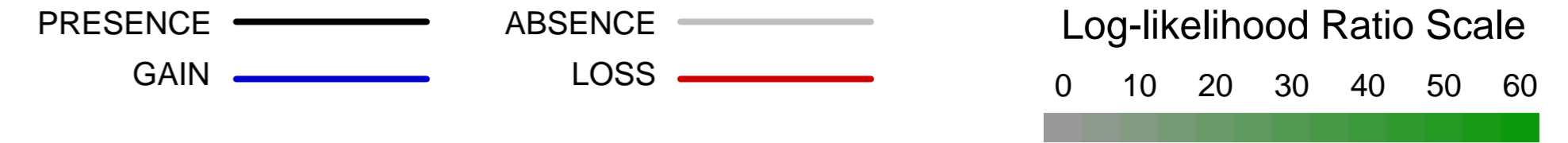


| PG | Protein | LLR | Notes |
|----|---------|-----|-------|
| A  | MLH1    |     | 1     |
| A  | MLH3    |     | 1     |
| A  | PMS2    |     | 1     |

1: Mismatch repair

ECM 3, Gene set "Mismatch repair", Page 1

Num of ECM Genes: 1. Num of Predicted Genes: 0



PG

Protein

LLR Notes

1: Mismatch repair

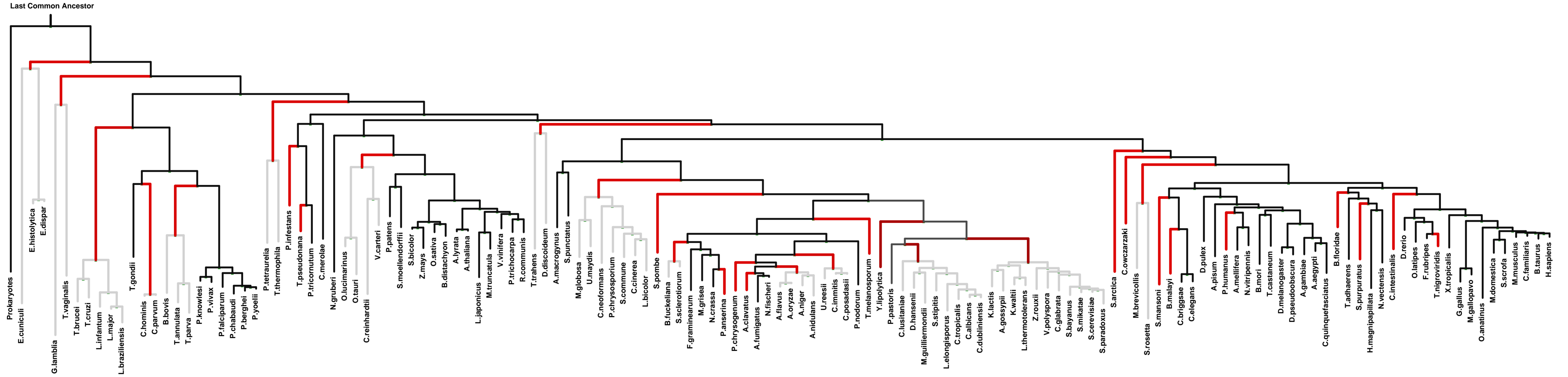
ECM 4, Gene set "Mismatch repair", Page 1

Num of ECM Genes: 1. Num of Predicted Genes: 25

PRESENCE ———— ABSENCE ————  
GAIN ———— LOSS ————

Log-likelihood Ratio Scale

0 10 20 30 40 50 60



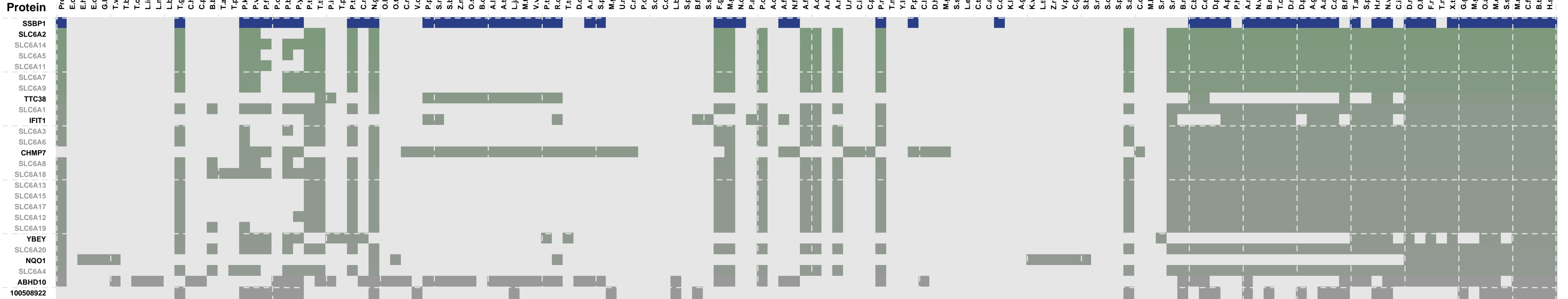
Protists

Plants

Fungi

Metazoa

PG



LLR Notes

11.5 1/2/3  
11.3  
10.7  
10.6  
9.3  
9.3  
7.3  
5.2  
5.0  
5.0  
4.9  
4.9  
4.9  
4.6  
4.0  
4.0  
4.0  
4.0  
4.0  
4.0  
3.7  
3.7  
3.3  
2.9  
2.4  
0.4  
0.3

1: Mismatch repair || 2: Homologous recombination || 3: DNA replication || 4: Endocytosis