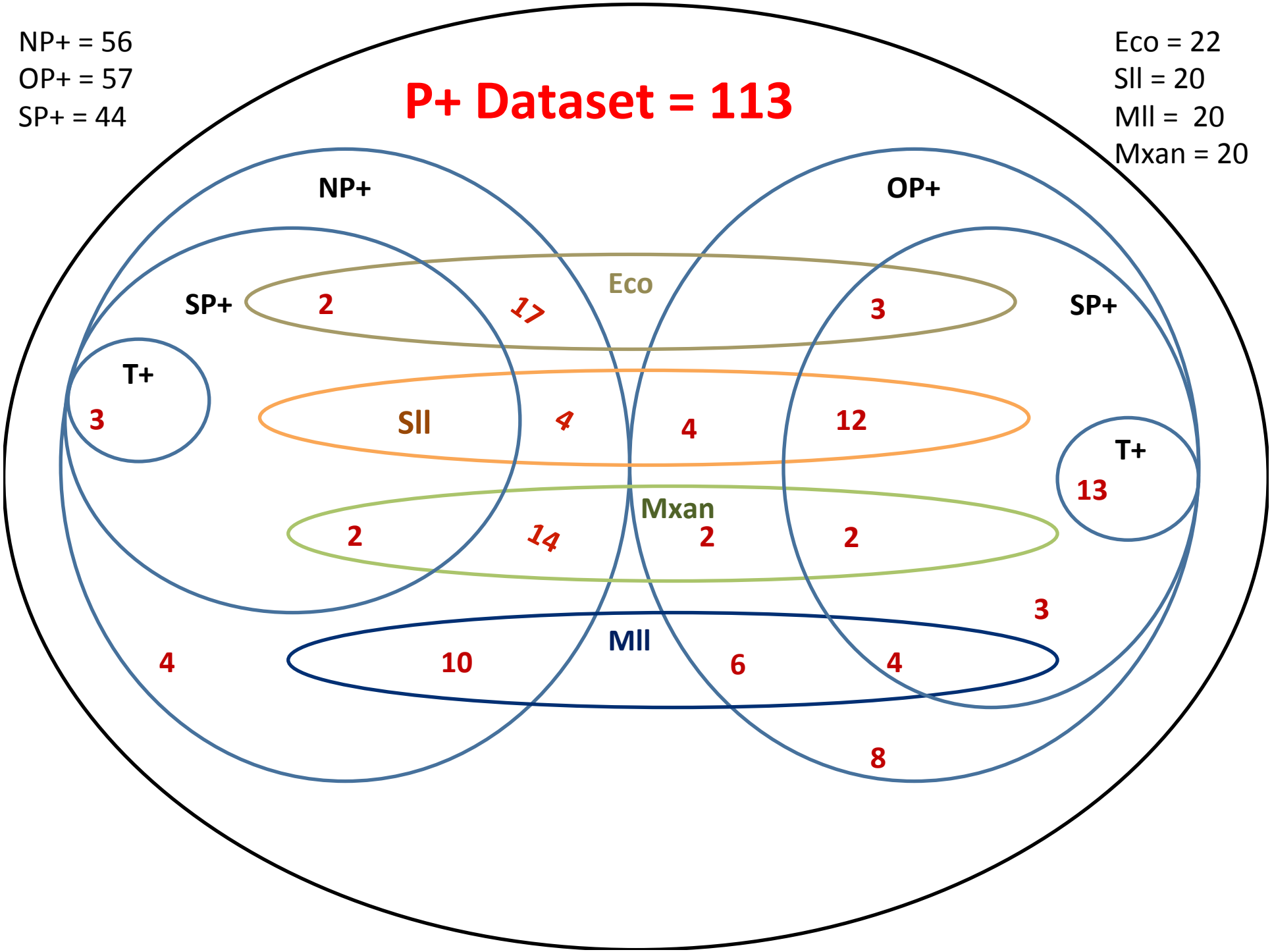
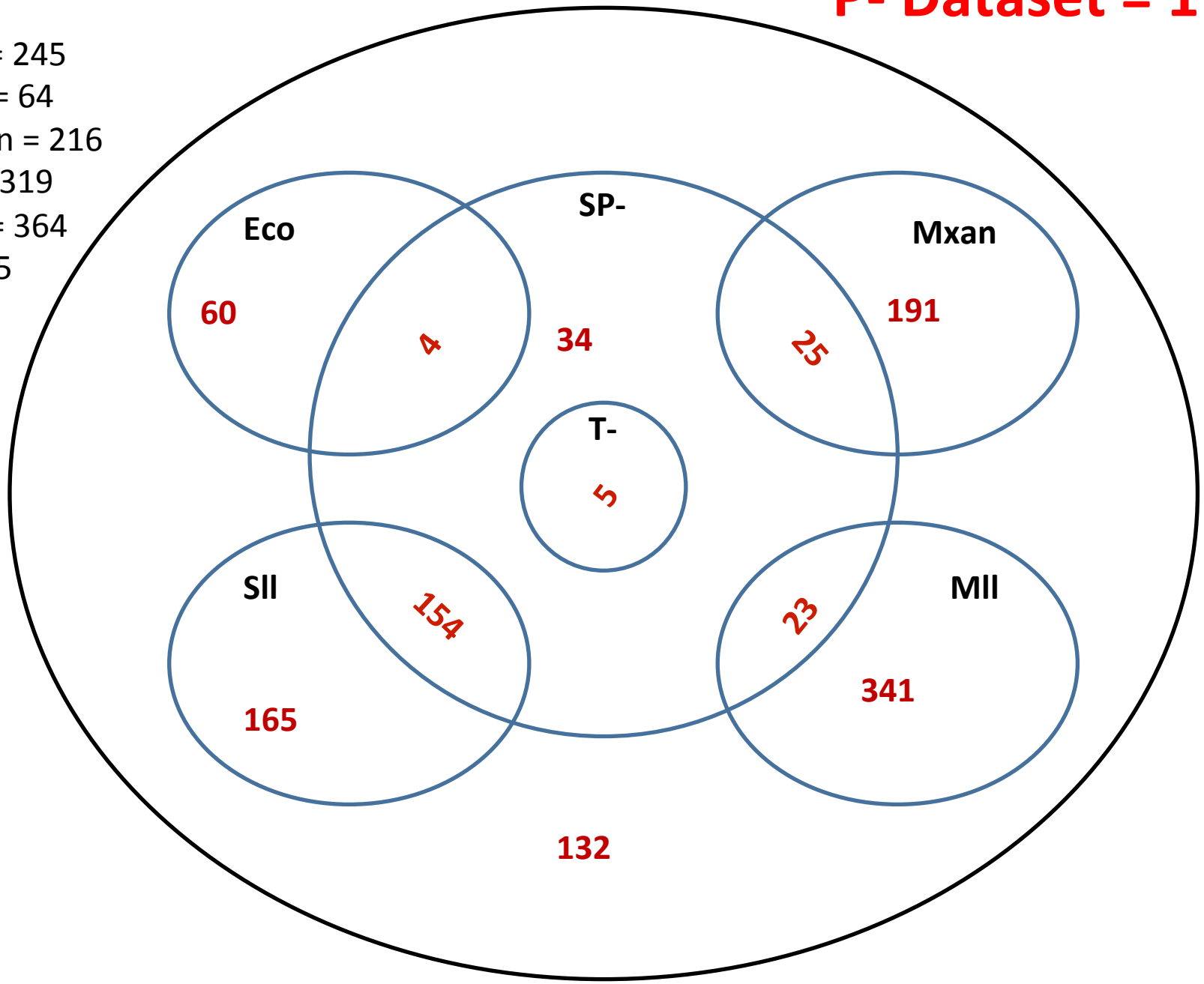
Acronyms and definitions

- **P+:** Gold standard dataset interacting
- **P-:** Gold standard dataset of non-interacting
- **NP+:** Subset of P+ composed of encoded by neighbouring genes
- **OP+:** Subset of P+ composed of encoded by non-neighbouring genes
- **T:** The dataset compiled in Burger and van Nimwegen, Mol Sys Biol 2008
- **T+:** Interacting included in T dataset
- **T-:** Non-interacting included T dataset
- **SP+:** Subset of P+ dataset composed of present in Burger and van Nimwegen's database (also used to compare to STRING database)
- **SP-:** Subset of P+ dataset composed of present in Burger and van Nimwegen's database (also used to compare to STRING database)
- **Eco =** *Escherichia coli* K-12 MG1655
- **Sll =** *Synechocystis sp*. PCC6803
- **Mll =** *Mesorhizobium loti* MAFF303099
- **Mxan =** *Myxococcus xanthus* DK 1622


NP+ = 56
OP+ = 57
SP+ = 44
P+ Dataset = 113
Eco = 22
Sll = 20
Mll = 20
Mxan = 20
NP+
OP+
Eco
SP+
SP+
T+
Sll
T+
Mxan
Mll
2
17
3
3
4
4
12
13
2
14
2
2
3
4
10
6
4
8

P- Dataset = 1134

SP- = 245
Eco = 64
Mxan = 216
Sll = 319
Mll = 364
T- = 5