Table 1

Number of genes products predicted to be in an RGC or predicted to participate on plant resistance processes using clustering, classification and influential analysis. E: Euclidean distance, M: Manhattan distance.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Method | NI:non-resistance | NI: resistance | I:non-resistance | I:resistance |
| K-means (E) | 4818 | 58 | 4681 | 73 |
| K-means (M) | 4760 | 57 | 4681 | 73 |
| AGNES (E) | 6463 | 87 | 6863 | 99 |
| AGNES (M) | 6598 | 86 | 6756 | 99 |
| DIANA (E) | 11128 | 145 | 12632 | 158 |
| DIANA (M) | 4760 | 57 | 12325 | 156 |
| Kohonen | 5120 | 105 | 4323 | 75 |
| SVM | 98 | 69 | 0 | 105 |
| Pearson | 36 | 0 | 0 | 0 |
| Leverage | 1 | 165 | 0 | 170 |
| Cook | 0 | 2 | 1 | 170 |
| DfBeta | 0 | 0 | 4 | 0 |
| DfBeta for Period | 17 | 0 | 0 | 10 |
| DfBeta for Resistance | 1 | 9 | 13 | 7 |
| DfBeta for Period\*Resistance | 1 | 8 | 0 | 0 |

Table 2

List of common genes predicted as resistance genes by clustering methods, SVM classification and influential analysis on GEE regression models.

|  |  |
| --- | --- |
| Gene ID  | Function |
| 1-1-2.1.13.17 | Avr9/Cf-9 rapidly elicited protein 146 *Nicotiana tabacum* |
| 1-1-2.4.18.3 | AvrPto-dependent Pto-interacting protein 3 *Lycopersicon esculentum* |
| 1-1-3.3.11.4 | WIZZ *Nicotiana tabacum* |
| 1-1-3.4.10.21 | Protein kinase-coding resistance protein *Nicotiana repanda* |
| 1-1-4.1.17.2 | Putative disease resistance protein RGA4, identical *Solanum bulbocastanum* |
| 1-1-4.3.20.3 | Disease resistance protein RGA2, putative *Ricinus communis* |
| 1-1-7.4.19.21 | WRKY transcription factor 26 *Populus tomentosa x P. bolleana* x *P. tomentosa* |
| 1-1-8.2.15.5 | Disease resistance protein RPS5, putative *Ricinus communis* |
| 1-1-8.2.16.9 | Avr9/Cf-9 induced kinase 1 *Nicotiana tabacum* |
| 1-1-8.4.11.17 | WRKY *Solanum lycopersicum* |