**Supplementary Material S3.** Selected 158 features and their occurrences in the 10 feature sets of the 10 datasets (√: feature is in the feature set; ☓: feature is not included in the feature set)

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Feature name** | **Frequency** | **Feature set on** | | | | | | | | | |
| ***DS*1** | ***DS*2** | ***DS*3** | ***DS*4** | ***DS*5** | ***DS*6** | ***DS*7** | ***DS*8** | ***DS*9** | ***DS*10** |
| hsa03010\_1+hsa03010\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0000184\_1+GO:0000184\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0000375\_1+GO:0000375\_2 | 7 | √ | √ | ☓ | √ | ☓ | √ | √ | √ | ☓ | √ |
| GO:0000956\_1+GO:0000956\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0003674\_1+GO:0003674\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0003676\_1+GO:0003676\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0003723\_1+GO:0003723\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0005488\_1+GO:0005488\_2 | 1 | ☓ | ☓ | ☓ | ☓ | ☓ | ☓ | ☓ | ☓ | √ | ☓ |
| GO:0005515\_1+GO:0005515\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0005622\_1+GO:0005622\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0005623\_1+GO:0005623\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0005634\_1+GO:0005634\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0005654\_1+GO:0005654\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0005694\_1+GO:0005694\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0005730\_1+GO:0005730\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0005829\_1+GO:0005829\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0006139\_1+GO:0006139\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0006259\_1+GO:0006259\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0006281\_1+GO:0006281\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0006351\_1+GO:0006351\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0006396\_1+GO:0006396\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0006401\_1+GO:0006401\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0006402\_1+GO:0006402\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0006412\_1+GO:0006412\_2 | 6 | √ | ☓ | √ | √ | ☓ | ☓ | √ | ☓ | √ | √ |
| GO:0006413\_1+GO:0006413\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0006414\_1+GO:0006414\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0006415\_1+GO:0006415\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0006613\_1+GO:0006613\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0006614\_1+GO:0006614\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0006725\_1+GO:0006725\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0006807\_1+GO:0006807\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0006974\_1+GO:0006974\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0006996\_1+GO:0006996\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0008150\_1+GO:0008150\_2 | 9 | √ | √ | √ | √ | √ | ☓ | √ | √ | √ | √ |
| GO:0009057\_1+GO:0009057\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0009058\_1+GO:0009058\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0009059\_1+GO:0009059\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0009889\_1+GO:0009889\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0009987\_1+GO:0009987\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0010467\_1+GO:0010467\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0010468\_1+GO:0010468\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0010556\_1+GO:0010556\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0016021\_1+GO:0016021\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0016032\_1+GO:0016032\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0016043\_1+GO:0016043\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0016070\_1+GO:0016070\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0016071\_1+GO:0016071\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0016604\_1+GO:0016604\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0018130\_1+GO:0018130\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0019058\_1+GO:0019058\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0019080\_1+GO:0019080\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0019083\_1+GO:0019083\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0019219\_1+GO:0019219\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0019222\_1+GO:0019222\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0019438\_1+GO:0019438\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0019439\_1+GO:0019439\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0019538\_1+GO:0019538\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0022411\_1+GO:0022411\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0022613\_1+GO:0022613\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0022626\_1+GO:0022626\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0030529\_1+GO:0030529\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0031224\_1+GO:0031224\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0031323\_1+GO:0031323\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0031326\_1+GO:0031326\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0031974\_1+GO:0031974\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0031981\_1+GO:0031981\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0032774\_1+GO:0032774\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0032984\_1+GO:0032984\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0032991\_1+GO:0032991\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0033554\_1+GO:0033554\_2 | 8 | √ | √ | √ | √ | ☓ | ☓ | √ | √ | √ | √ |
| GO:0034641\_1+GO:0034641\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0034645\_1+GO:0034645\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0034654\_1+GO:0034654\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0034655\_1+GO:0034655\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0042254\_1+GO:0042254\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0043170\_1+GO:0043170\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0043228\_1+GO:0043228\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0043232\_1+GO:0043232\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0043233\_1+GO:0043233\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0043234\_1+GO:0043234\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0043241\_1+GO:0043241\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0043624\_1+GO:0043624\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0043933\_1+GO:0043933\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0044033\_1+GO:0044033\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0044237\_1+GO:0044237\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0044238\_1+GO:0044238\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0044249\_1+GO:0044249\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0044260\_1+GO:0044260\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0044265\_1+GO:0044265\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0044267\_1+GO:0044267\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0044270\_1+GO:0044270\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0044271\_1+GO:0044271\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0044391\_1+GO:0044391\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0044403\_1+GO:0044403\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0044419\_1+GO:0044419\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0044422\_1+GO:0044422\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0044424\_1+GO:0044424\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0044427\_1+GO:0044427\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0044428\_1+GO:0044428\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0044445\_1+GO:0044445\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0044446\_1+GO:0044446\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0044451\_1+GO:0044451\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0044464\_1+GO:0044464\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0044710\_1+GO:0044710\_2 | 5 | ☓ | √ | √ | √ | ☓ | ☓ | ☓ | √ | ☓ | √ |
| GO:0044764\_1+GO:0044764\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0044822\_1+GO:0044822\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0045047\_1+GO:0045047\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0046483\_1+GO:0046483\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0046700\_1+GO:0046700\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0051171\_1+GO:0051171\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0051252\_1+GO:0051252\_2 | 3 | ☓ | ☓ | ☓ | √ | √ | √ | ☓ | ☓ | ☓ | ☓ |
| GO:0051276\_1+GO:0051276\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0060255\_1+GO:0060255\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0070013\_1+GO:0070013\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0070972\_1+GO:0070972\_2 | 9 | √ | √ | √ | √ | √ | √ | √ | ☓ | √ | √ |
| GO:0071704\_1+GO:0071704\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0071840\_1+GO:0071840\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0072599\_1+GO:0072599\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0080090\_1+GO:0080090\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0090304\_1+GO:0090304\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:0097159\_1+GO:0097159\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:1901360\_1+GO:1901360\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:1901361\_1+GO:1901361\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:1901362\_1+GO:1901362\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:1901363\_1+GO:1901363\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:1901576\_1+GO:1901576\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:1990234\_1+GO:1990234\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| GO:2000112\_1+GO:2000112\_2 | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| abs(GO:0003674\_1-GO:0003674\_2) | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| abs(GO:0005488\_1-GO:0005488\_2) | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| abs(GO:0005622\_1-GO:0005622\_2) | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| abs(GO:0005623\_1-GO:0005623\_2) | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| abs(GO:0005634\_1-GO:0005634\_2) | 1 | ☓ | ☓ | ☓ | ☓ | ☓ | ☓ | √ | ☓ | ☓ | ☓ |
| abs(GO:0005730\_1-GO:0005730\_2) | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| abs(GO:0008152\_1-GO:0008152\_2) | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| abs(GO:0009987\_1-GO:0009987\_2) | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| abs(GO:0016020\_1-GO:0016020\_2) | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| abs(GO:0016021\_1-GO:0016021\_2) | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| abs(GO:0016032\_1-GO:0016032\_2) | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| abs(GO:0019083\_1-GO:0019083\_2) | 6 | √ | √ | ☓ | √ | ☓ | ☓ | √ | ☓ | √ | √ |
| abs(GO:0022613\_1-GO:0022613\_2) | 6 | √ | √ | √ | ☓ | ☓ | ☓ | √ | √ | √ | ☓ |
| abs(GO:0031224\_1-GO:0031224\_2) | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| abs(GO:0043170\_1-GO:0043170\_2) | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| abs(GO:0043226\_1-GO:0043226\_2) | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| abs(GO:0043227\_1-GO:0043227\_2) | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| abs(GO:0043229\_1-GO:0043229\_2) | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| abs(GO:0043231\_1-GO:0043231\_2) | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| abs(GO:0044237\_1-GO:0044237\_2) | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| abs(GO:0044238\_1-GO:0044238\_2) | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| abs(GO:0044260\_1-GO:0044260\_2) | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| abs(GO:0044403\_1-GO:0044403\_2) | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| abs(GO:0044419\_1-GO:0044419\_2) | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| abs(GO:0044424\_1-GO:0044424\_2) | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| abs(GO:0044425\_1-GO:0044425\_2) | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| abs(GO:0044451\_1-GO:0044451\_2) | 7 | √ | √ | √ | √ | ☓ | √ | √ | √ | ☓ | ☓ |
| abs(GO:0044464\_1-GO:0044464\_2) | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| abs(GO:0044764\_1-GO:0044764\_2) | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| abs(GO:0071704\_1-GO:0071704\_2) | 10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |