Supporting Information

Table of contents

Supporting Texts

Supporting Tables

Supporting Figures

Figure S1: The performance of six expression profile-based methods on the test datasets for 8 species.

Figure S2: The precision-recall curves of six expression profile-based methods on the test datasets for 8 species.

Figure S3: The AVG_WFS values of six measures for three GO aspects in 8 individual species.

Figure S4: The scattering plots of weights versus F1-scores of 100 templates for the gene MIRLET7C over TNP, MR, and PCC.

Figure S5: The $Fmax$ values of nine GO prediction methods on the test datasets for 8 species.

Figure S6: The $AUPR$ values of nine GO prediction methods on the test datasets for 8 species.

Figure S7: The precision-recall curves of five GO prediction methods on the test datasets for 8 species.

Figure S8: Comparison of mean and median AUROC values of three GO aspects by different methods on the common dataset.

Figure S9: Comparison of $Fmax$ and $AUPR$ values of three GO aspects by different methods on the common dataset.

Figure S10: The distribution of sequence identities for 10000 gene-gene pairs and 10000 mapped protein-protein pairs.

Supporting Texts

Text S1. The procedures of GBA strategy for expression profile-based GO prediction

In the guilty-by-association (GBA) strategy, we select the template genes which have the highest similarity with query gene in terms of expression profiles, and then use the GO terms of templates to annotate the query, as follows.

Training stage

In a training dataset, the expression profiles of all genes can be represented as a matrix $\boldsymbol{E} = (e_{ij})_{m\times l}$, where the i-th row of $\boldsymbol{E}$ is the expression profile for the i-th gene and denoted as $\boldsymbol{e}_i = (e_{i1}, e_{i2}, \dots, e_{il})^T$, m is the total number of training genes, l is the number of experimental samples in microarray technology [1], and e_{ij} is the expression value of the i-th gene on the j-th sample. We orderly execute z-score normalization [2] and principal component analysis (PCA) [3] on expression profile matrix $\boldsymbol{E}$ to obtain a normalized matrix $\boldsymbol{E}^n = (e_{ij}^n)_{m\times h}$, where the i-th row of $\boldsymbol{E}^n$, denoted as $\boldsymbol{e}_i^n = (e_{i1}^n, e_{i2}^n, \dots, e_{ih}^n)^T$, is the normalized expression profile vector for the i-th training gene.

Prediction stage

For a query gene, its expression profile can be represented as a vector $\boldsymbol{e}^q = (e_1^q, e_2^q, \dots, e_l^q)^T$. First, the z-score normalization and PCA are orderly executed on the expression profile vector $\boldsymbol{e}^q$ to obtain a normalized vector $\boldsymbol{e}_q^n = \left(e_1^{nq}, e_2^{nq}, \dots, e_h^{nq}\right)^T$. Then, for each training gene i, we calculate its similarity score with query based on the normalized vector $\boldsymbol{e}_q^n$ and $\boldsymbol{e}_i^n$. Next, we rank m training genes based on the similarity scores in descending order. Finally, we select the top K training genes as templates to annotate the GO terms of query. Specifically, the confidence score that the query is associated with GO term Q_j can be calculated as follows:

$$S(Q_j)_{GBA} = \frac{\sum_{k=1}^{K} w_k \cdot I_k(Q_j)}{\sum_{k=1}^{K} w_k} \tag{S1}$$

$$w_k = 1 - (r_k - 1)/K \tag{S2}$$

where w_k is the weight for the k-th template, and r_k is the rank of the k-th template; $I_k(Q_j) = 1$, if the k-th template is associated with Q_j in the experimental annotation;

otherwise, $I_k(Q_j) = 0$.

In this work, the similarity score of expression profiles between two genes are measured by four unsupervised methods, including Pearson correlation coefficient (PCC) [4], Spearman rank correlation (SRC) [5], mutual rank (MR) [6], and Euclidean distance (ED) [7], and a recently proposed supervised method, i.e., metric learning for co-expression (MLC) [8].

The PCC between the i-th training gene and query gene is calculated as follows:

$$\mathrm{PCC}(\boldsymbol{e}_i^n, \boldsymbol{e}_q^n) = \frac{\sum_{j=1}^{h}(e_{ij}^n - \overline{\boldsymbol{e}_i^n}) \cdot (e_j^{nq} - \overline{\boldsymbol{e}_q^n})}{\sqrt{\sum_{i=1}^{h}(e_{ij}^n - \overline{\boldsymbol{e}_i^n})^2} \cdot \sqrt{\sum_{i=1}^{h}(e_j^{nq} - \overline{\boldsymbol{e}_q^n})^2}} \tag{S3}$$

where $\overline{\boldsymbol{e}_i^n}$ and $\overline{\boldsymbol{e}_q^n}$ are mean values for $\boldsymbol{e}_i^n$ and $\boldsymbol{e}_q^n$, respectively.

The SRC between the i-th training gene and query gene is calculated as follows:

$$\mathrm{SRC}(\boldsymbol{e}_i^n, \boldsymbol{e}_q^n) = 1 - \frac{6\sum_{j=1}^{h}(r_{ij} - r_j)^2}{h(h^2-1)} \tag{S4}$$

where r_{ij} is rank of e_{ij}^n in the elements of $\boldsymbol{e}_i^n$ in ascending order, r_j is the rank of e_j^{nq} in the elements of $\boldsymbol{e}_q^n$ in ascending order.

Due to the long computation time of MR values, we directly download MR values of genes from COXPRESdb [9] and ATTED-II databases [6]. In a species with M genes, the MR value between gene i and gene j is calculate as follows. First, we calculate the PCC values between gene i and the remaining $M-1$ genes based on the corresponding expression profile vectors, and rank the $M-1$ genes based on the PCC values in descending order. Similarly, we calculate the PCC values between gene j and the remaining $M-1$ genes, and rank the $M-1$ genes in descending order based on PCC values. Then, the MR value between genes i and j can be calculated:

$$MR(i,j) = \sqrt{rank(i) \cdot rank(j)} \tag{S5}$$

where $rank(i)$ is the rank of gene i in $M-1$ genes for gene j, and $rank(j)$ is the rank of gene j in $M-1$ genes for gene i.

The ED between the i-th training gene and query gene is calculated as follows:

$$\mathrm{ED}(\boldsymbol{e}_i^n, \boldsymbol{e}_q^n) = \sqrt{\sum_{j=1}^{h}\left(e_{ij}^n - e_j^{nq}\right)^2} \tag{S6}$$

In MLC, the similarity between the i-th training gene and query gene is measured by weight inner product (WIP) as follows:

$$\mathrm{WIP}(\boldsymbol{e}_i^n, \boldsymbol{e}_q^n) = (\boldsymbol{e}_i^n)^T \cdot W \cdot \boldsymbol{e}_q^n \tag{S7}$$

where $W = \mathrm{diag}(w)$ is a diagonal matrix and can be optimized by the Broyden-

Fletcher-Goldfarb-Shanno method [10].

The higher values of PCC, SRC, and WIP indicate the higher similarity, while the lower values of MR and ED mean the higher similarity.

Text S2. The performances of six expression profile-based GO prediction methods for each individual species

For each of 8 species, we will evaluate the performances of six expression profile-based GO prediction methods on the corresponding test dataset. For each method, we execute it 10 times and then use the average of all prediction results as the final result.

Figure S1 show the values of $Fmax$ and $AUPR$ for eight species via six expression profile-based methods. Table S2 summarizes the p-values of $Fmax$ and $AUPR$ values between TNP and other five methods in student's t-test [11] for 8 species. In comparison between TNP and MLC, we use two samples t-test [12] to calculate p-value due to that the prediction results in 10 times are different for MLC/TNP. In comparison between TNP and PCC, MR, SRC, ED, we use single samples t-test [13] to calculate p-value, because the prediction results in 10 times are same for PCC/MR/SRC/ED. From Figure S1 and Table S2, we can observe that TNP achieves the highest values of $Fmax$ and $AUPR$ among six methods for each GO aspect in each species. For example, in human species, the improvements of $Fmax$ values between TNP and MR are 12.7%, 8.2%, 3.8%, respectively, with p-values of 1.29×10^{-04}, 7.83×10^{-09}, and 5.07×10^{-07} for MF, BP, and CC aspects. As another example, the average improvement of $AUPR$ values of three GO aspects between TNP and the second best performer is 8.6% with p-values<0.05 for arabidopsis species.

Figure S2 plots the precision-recall (PR) curves of six expression-profile based methods for three GO aspects in 8 species. For each GO aspect in each species, we can find that TNP has the highest precision values among six expression profile-based methods at all different recall rates.

Text S3. The functional similarity for genes

The functional similarity of two genes is measured by the F1-score between their experimental GO terms. For a gene pair (i, j), the F1-score between their GO terms is defined as:

$$\mathrm{F1-score} = 2(\mathrm{pre}\times\mathrm{rec})/(\mathrm{pre}+\mathrm{rec}),\ \mathrm{pre} = ns/n_1,\ \mathrm{rec} = ns/n_2 \quad \text{(S8)}$$

where ns is the number of same GO terms between two genes, n_1 and n_2 are the

numbers of GO terms for genes i and j, respectively.

Text S4. The performances of nine GO prediction methods for each individual species

For each of 8 species, we will compare the performances of four individual methods (i.e., EPGP, GSAGP, PSAGP, and NGP) and five combination methods (i.e., GPN, EPN, EGN, EGP, and EGPN) on the corresponding test dataset. For each combination method, we execute it 10 times and then use the average of all prediction results as the final result.

Figures S5-S6 illustrate the values of $Fmax$ and $AUPR$ for nine GO prediction methods in 8 species. Table S4 show the p-values of $Fmax$ and $AUPR$ values between EGPN and other eight methods in student's t-test [11] for 8 species. In comparison between EGPN and four combination methods (i.e., GPN, EPN, EGN, and EGP), we use two samples t-test [12] to calculate p-value due to that the prediction results in 10 times are different for them. In comparison between EGPN and four individual methods (i.e., EPGP, GSAGP, PSAGP, and NGP), we use single samples t-test [13] to calculate p-value. From Figures S5-S6 and Table S4, we can find the $Fmax$ and $AUPR$ values of EGPN are much higher than that of four individual methods for each species. Moreover, from the view of $Fmax$, in species of arabidopsis and fly, EGPN achieves the better performance than other four combination methods for each GO aspect; in species of human, mouse, rat and nematoda, EGPN occupies one of the top two positions among five combination methods for each GO aspect; as for the remaining two species, EGPN shows the best performance in MF/BP for budding yeast and BP/CC for fission yeast. These observations further demonstrate that each individual method contributes to improving prediction performance.

Figure S7 plot the precision-recall (PR) curves of four individual methods and EPGN for three GO aspects in 8 species. For each GO aspect in each species, we can observe that the PR curve of EGPN is continuously higher than that of four individual methods.

Text S5. Finding common genes and GO terms between our datasets and GENETICA's datasets

In the web page (http://genetica-network.com), GENETICA provides the prediction scores and real labels of three GO aspects for 19,635 genes in human species and 18425

genes in mouse species. Specifically, in human species, each gene is associated with the prediction scores and real labels of 843 MF, 4203 BP, and 528 CC GO terms. As for mouse species, each gene is associated with the prediction scores and real labels of 833 MF, 4188 BP, and 525 CC terms.

In GENETICA's datasets, 879 genes, 1176 genes, and 1241 genes for MF, BP, and CC aspects, respectively, can be found in our test dataset for human species; 572 genes, 880 genes, and 737 genes for three GO aspects can be separately found in our test dataset for mouse species. The GO terms in GENETICA and our work are represented as GO names and GO IDs, respectively. Due to the different versions of gene ontology databases, only 738 MF, 3980 BP and 476 CC GO names in GENETICA's can be correctly mapped as the corresponding GO IDs in our dataset for human species; As for mouse species, there are 727 MF, 3965 BP, and 474 CC terms in common between our work and GENETICA. Moreover, we only consider the GO terms whose frequencies are more than 20 both in our training datasets and GENETICA's datasets. After this, there are 879 genes annotated with 287 MF terms, 1176 genes with 1340 BP terms, and 1241 genes with 186 CC terms for human species in common between our test dataset and GENETICA's dataset; As for mouse species, there are 572 genes with 149 MF terms, 880 genes with 1230 BP terms, and 737 genes with 128 CC terms in common. For each GO term Q_i, all genes are assigned with the prediction scores and real labels. Specifically, if a gene is associated with Q_i both in our test dataset and GENETICA's dataset, we label it as "1"; otherwise, it is labeled as "0".

Text S6. Finding common genes and GO terms between our datasets and GeneNetwork's datasets

In the web page (https://www.genenetwork.nl/), we can use command "GET https://www.genenetwork.nl/api/v1/gene/geneName?db=database" to download the GO information file generated by GeneNetwork for each query gene in each GO aspect. The information file contains all GO terms in the experimental function annotation and 100 predicted GO terms with scores for a query. GeneNetwork provides all of the information files in three GO aspects for 56435 genes.

In GeneNetwork's dataset, 918 genes, 1230 genes, and 1328 genes for MF, BP and CC aspects, respectively, can be found in our test dataset for human species. Moreover, there are 655 MF, 2776 BP, and 536 CC terms in common between our work and GeneNetwork. In this work, we only consider the GO terms whose frequencies are more

than 20 both in our training dataset and GeneNetwork's dataset. After this, there are 918 genes associated with 165 MF terms, 1230 genes with 522 BP terms, and 1328 genes with 182 CC terms, in common, for human species between our test dataset and GeneNetwork's dataset. For each GO term Q_i, all genes are assigned with the prediction scores and real labels. Specifically, if a gene is associated with Q_i both in our test dataset and GeneNetwork's dataset, we label it as "1"; otherwise, it is labeled as "0".

Text S7. Comparison with the existing gene function prediction models in gene-center level.

We further compared our methods (TNP and TripletGO) with the existing gene function predictors (GENETICA and GeneNetwork) in gene-center level. In GENETICA and GeneNetwork, the numerical distributions of prediction scores of genes are different for each GO aspect. For example, in MF aspect, the prediction scores of GO terms for all human genes by GENETICA are range from 0.618167 to 15.79360; in BP aspect, the prediction scores by GENETICA are range from 0.61976 to 19.75050. Therefore, we firstly normalize the prediction scores in the range of 0 to 1 for each GO aspect using Min-Max Normalization. Specifically, for each GO aspect, the prediction score S_i is normalized as:

$$S_i^n = \frac{S_i - S_{min}}{S_{max} - S_{min}} \quad \text{(S9)}$$

where S_{max} and S_{min} are the max and min values, respectively, in all prediction scores.

Based on the normalized prediction scores, the prediction performances for GENETICA and GeneNetwork can be transformed from term-center metric (AUROC) to gene-center metric ($Fmax$ and $AUPR$). Figure S9 (A-B) shows the $Fmax$ and $AUPR$ values of TNP, TripletGO and GENETICA in human species (879 genes with 287 MF terms, 1176 genes with 1340 BP terms, and 1241 genes with 186 CC terms) and mouse species (572 genes with 149 MF terms, 880 genes with 1230 BP terms, and 737 genes with 128 CC terms). Moreover, we further compared our methods with GeneNetwork in human species (918 genes with 165 MF terms, 1230 genes with 522 BP terms, and 1328 genes with 182 CC terms), as shown in Figure S9 (C). From Figure S9 (A-C), we find that the proposed TNP and TripletGO show the significantly better performance than GENETICA and GeneNetwork for each GO aspect. For example,

from the view of $Fmax$, TNP achieves the improvements of 112.8%, 60.7% and 59.1%, respectively, for MF, BP, and CC aspects of human species in comparison with GENETICA. It cannot escape our notice that the $Fmax$ and $AUPR$ values of TNP and TripletGO in Figure S9 are significantly lower than the corresponding values in the previous figures (Figures S1, S5 and S6), especially for CC aspect. The reason can be explained as follows. Frist, we remove a part of genes and terms, which are not included in the GENETICA's dataset, from our test dataset. Second, the GO annotations of genes for our work and GENETICA are originated from different databases. Specifically, the GO annotation of genes in our work are downloaded from NCBI with the version of "2021-02-23"; in GENETICA, the GO annotations are extracted from Broad Institute Molecular Signatures Database v6.2. Therefore, the GO annotations for a part of test genes are different between our work and GENETICA. For example, for gene ARHGAP1 (Entrez ID: 392), we listed the corresponding GO annotations for CC aspect in different works as follows.

(1) Our work: GO:0005768, GO:0110165, GO:0016020, GO:0043227, GO:0043226, GO:0097708, GO:0005737, GO:0031982, GO:0097443, GO:0031410, GO:0010008, GO:0098588, GO:0031090, GO:0048471, GO:0005829 (15 terms).

(2) GENETICA: GO:0048471, GO:0110165 (2 terms)

(3) GeneNetwork: GO:0005737, GO:0016020, GO:0043230, GO:0070062, GO:0005829, GO:0098588, GO:0010008, GO:0031090, GO:0048471, GO:1903561, GO:0043227, GO:0043226, GO:0110165, GO:0031982 (14 terms)

(4) Common terms between our work and GENETICA: GO:0048471, GO:0110165 (2 terms)

(5) Common terms between our work and GeneNetwork: GO:0005737, GO:0031090, GO:0005829, GO:0016020, GO:0031982, GO:0043226, GO:0048471, GO:0043227, GO:0098588, GO:0110165, GO:0010008 (11 terms).

We can notice that there are only 2 common GO terms (GO:0048471, GO:0110165) for the annotation of gene ARHGAP1 between our work and GENETICA. To compare our method and GENETICA in fairness, these 2 common GO terms are used as "gold standard" for GO annotation of gene ARHGAP1 and the remaining terms are ignored. Therefore, some predicted terms, which are considered as true positives in previous Figures (Figures S1, S5 and S6), are viewed as false positives in the comparison with GENETICA, which further leads the significant performance degradation of TNP and TripletGO in this section.

Text S8. The predicted GO terms of ten GO prediction methods for genes GALNT4 and MIRLET7C

GALNT4

MR	**GO:0043231 GO:0016020 GO:0005622 GO:0043229 GO:0043227 GO:0043226 GO:0110165 GO:0031090** GO:0005654 GO:0005829 GO:0032991
PCC	**GO:0043231 GO:0016020 GO:0005622 GO:0043229 GO:0043227 GO:0043226 GO:0110165 GO:0031090** GO:0005829
MLC	**GO:0043231 GO:0016020 GO:0005622 GO:0043229 GO:0043227 GO:0043226 GO:0110165** GO:0005829 GO:0032991
SRC	**GO:0043231 GO:0016020 GO:0005622 GO:0043229 GO:0043227 GO:0043226 GO:0110165** GO:0005654 GO:0005829 GO:0005886
ED	**GO:0043231 GO:0016020 GO:0005622 GO:0043229 GO:0043227 GO:0043226 GO:0110165 GO:0031090** GO:0005829
TNP	**GO:0043231 GO:0016020 GO:0005622 GO:0043229 GO:0043227 GO:0043226 GO:0110165 GO:0031090**
GSAGP	**GO:0043231 GO:0016020 GO:0005622 GO:0043229 GO:0043227 GO:0043226 GO:0110165** GO:0005654 GO:0005829 GO:0005886
PSAGP	**GO:0031090 GO:0098588 GO:0016020 GO:0000139 GO:0043227 GO:0043226 GO:0110165** GO:0031410 GO:0030133 GO:0097708 GO:0031982
NGP	**GO:0043231 GO:0016020 GO:0005622 GO:0043229 GO:0043227 GO:0043226 GO:0110165** GO:0005829 GO:0032991 GO:0005634 GO:0005886
TripletGO	**GO:0031090 GO:0098588 GO:0016020 GO:0005622 GO:0043229 GO:0000139 GO:0043227 GO:0043226 GO:0110165 GO:0043231**
Annotation	**GO:0005794 GO:0043231 GO:0098588 GO:0016020 GO:0005622 GO:0043229 GO:0048471 GO:0000139 GO:0043227 GO:0043226 GO:0110165 GO:0031090**

MIRLET7C:

MR	**GO:0110165** GO:0016020 GO:0005886
PCC	**GO:0110165** GO:0016020 GO:0005886
MLC	**GO:0043227 GO:0043226 GO:0110165** GO:0016020 GO:0005886
SRC	**GO:0043231 GO:0005622 GO:0043229 GO:0043227 GO:0043226 GO:0110165** GO:0016020 GO:0005886
ED	**GO:0110165** GO:0016020 GO:0005886
TNP	**GO:0043231 GO:0005634 GO:0005622 GO:0043229 GO:0043227 GO:0043226 GO:0110165** GO:0005654
GSAGP	**GO:0005737 GO:0043231 GO:0005634 GO:0005622 GO:0043229 GO:0043227 GO:0043226 GO:0110165**
PSAGP	
NGP	**GO:0043231 GO:0005622 GO:0043229 GO:0043227 GO:0043226 GO:0110165** GO:0016020 GO:0032991
TripletGO	**GO:0005737 GO:0043231 GO:0005634 GO:0005622 GO:0043229 GO:0043227 GO:0043226 GO:0110165**
Annotation	**GO:0005737 GO:0043230 GO:0043231 GO:0070062 GO:0005634 GO:0005622 GO:0043229 GO:1903561 GO:0043227 GO:0043226 GO:0110165 GO:0031982**

Text S9. The construction procedures of Gene-GOA

First, we download all genes with GO annotation from National Center for

Biotechnology Information [14] (NCBI). Following the CAFA experiments [15, 16], we only select the genes annotated by at least one of the eight experimental evidence codes, including EXP, IDA, IPI, IMP, IGI, IEP, TAS, and IC. Moreover, to explicitly consider the hierarchical structure of GO terms, if a child term is annotated to a gene, all its direct and indirect parents, as defined by the "is_a" relation in gene ontology database [17] (http://geneontology.org/), are also annotated. The numbers of genes annotated with GO terms for MF, BP, and CC are 40160, 63543, and 55448, respectively, in Gene-GOA.

Text S10. Distance rank-based strategy

The distance rank-based strategy (DRBS) is executed on normalized embedding matrix of training genes ($\boldsymbol{U}^n$) and normalized embedding vector of query gene ($\boldsymbol{u}^q$) to obtain a confidence score vector, denoted as $\boldsymbol{s}^t = (s_1^t, s_2^t, \dots, s_r^t)^T$, where s_i^t is the confidence score that query is associated with the i-th GO term from the view of distance rank in the embedding space.

The details of DRBS are described as follows. First, we rank m training genes based the distances between the training and query genes in embedding space in ascending order. The distance between the i-th training gene and query gene is calculated as follows:

$$d(i, query) = \sum_{k=1}^{d_N} \left(u_{ik}^n - u_k^q\right)^2 / 4 \tag{S10}$$

Then, we select top K training genes which have the shortest distance with query in embedding space as templates to calculate the confidence scores of GO terms for query as follows:

$$s_i^t = \frac{\sum_{k=1}^{K} w_k \cdot I_k(i)}{\sum_{k=1}^{K} w_k} \tag{S11}$$

$$w_k = 1 - (r_k - 1)/K \tag{S12}$$

where w_k is the weight for the k-th template, and r_k is the rank of the k-th tempalte; $I_k(i) = 1$, if the k-th template is associated with the j-th GO term in the experimental function annotation; otherwise, $I_k(i) = 0$. In this work, the value of K is set to be 100.

Text S11. The construction procedures of genetic sequence database with GO annotation

To construct a genetic sequence database with GO annotation (GSD-GOA), the RNA sequences of all genes in Gene-GOA are extracted from NCBI [14]. If there is no

available RNA sequence for a gene, its genomic DNA sequence is selected. In addition, we discard a few genes which have no available RNA or genomic DNA sequences in NCBI. After this, GSD-GOA includes 39179 sequences with MF terms, 61699 sequences with BP terms, and 54117 sequences with CC terms.

Text S12. The construction procedures of protein sequence database

The protein sequence database (PSD) is constructed as follows. For each gene in Gene-GOA, we map it as the corresponding coding protein sequences in UniProt database [18] using a gene-protein mapping table. After this, 78170 genes can be mapped as 119876 protein sequences.

Text S13. The relationship between protein sequence identity and genetic sequence identity

In this work, we use three machine learning models, including liner regression (LR) [19], support vector regression (SVR) [20], and neural network with one hidden layer (NN), to fit the relationship between protein sequence identity and genetic sequence identity, as follows:

$$y = f(x)_{\mathrm{Model}} \tag{S13}$$

where x is the protein sequence identity, y is the genetic sequence identity, $\mathrm{Model} \in \{\mathrm{LR}, \mathrm{SVR}, \mathrm{NN}\}$.

First, we randomly select 100,000,000 gene-gene sequence pairs from NCBI, and map each gene-gene sequence pair as a protein-protein sequence pair in UniProt database by gene-protein mapping table. Then, to reduce computation time, we remove the protein sequences and gene sequences whose lengths are more than 10000. After this, the number of remaining gene-gene pairs (or the corresponding protein-protein pairs) is 974447. Next, we use standard Needleman-Wunsch algorithm [21] to calculate the sequence identity for each gene-gene pair and the corresponding protein-protein pair. Each gene-gene pair and the corresponding protein-protein pair are combined as a machine learning sample. More specifically, the sequence identity of protein-protein pair is used as the feature of sample, i.e., the input of machine learning model, and the sequence identity of gene-gene pair is used as the regression value of sample, i.e., the output of machine learning model. Finally, we separately train LR, SVR and NN on the 974447 samples, and found that $f(30\%)_{LR} = 58.6\%$, $f(30\%)_{SVM} = 59.3\%$, $f(30\%)_{NN} = 60.1\%$. In addition, we randomly select 10,000 gene-gene pairs and the

mapped protein-protein pairs, and then plot the corresponding sequence identities, as shown in Figure S10. For the protein-protein pairs whose sequence identities are equal to 30%, the sequence identities for the corresponding gene-gene pairs are near to 60%. In light of the above, we use 60% sequence identity as the cut-off to remove the homologous gene templates.

Supporting Tables

Table S1 The p-values between TNP and other five expression profile-based methods for $WAFmax$ and $WAAUPR$

Measure	GO aspect	(TNP, MR)	(TNP, PCC)	(TNP, MLC)	(TNP, SRC)	(TNP, ED)
$WAFmax$	MF	8.60×10^{-08}	1.70×10^{-10}	6.56×10^{-09}	3.19×10^{-09}	8.25×10^{-11}
	BP	6.11×10^{-11}	1.33×10^{-13}	1.71×10^{-12}	5.65×10^{-13}	8.64×10^{-14}
	CC	9.24×10^{-09}	3.67×10^{-11}	1.21×10^{-09}	3.85×10^{-10}	5.60×10^{-11}
$WAAUPR$	MF	5.13×10^{-10}	1.39×10^{-14}	1.14×10^{-11}	3.25×10^{-13}	1.04×10^{-14}
	BP	3.04×10^{-09}	4.43×10^{-13}	7.37×10^{-12}	2.61×10^{-12}	6.69×10^{-13}
	CC	1.82×10^{-13}	3.54×10^{-15}	1.00×10^{-13}	1.00×10^{-13}	6.58×10^{-15}

Table S2 The p-values between TNP and other five expression profile-based methods for $Fmax$ and $AUPR$ on 8 species

Species	Measure	GO aspect	(TNP, MR)	(TNP, PCC)	(TNP, MLC)	(TNP, SRC)	(TNP, ED)
Human	$Fmax$	MF	1.29×10^{-04}	3.08×10^{-08}	2.09×10^{-12}	1.80×10^{-05}	7.39×10^{-08}
		BP	7.83×10^{-09}	2.32×10^{-11}	7.00×10^{-18}	2.00×10^{-10}	3.23×10^{-11}
		CC	5.07×10^{-07}	1.95×10^{-10}	8.01×10^{-09}	2.53×10^{-09}	4.75×10^{-10}
	$AUPR$	MF	2.42×10^{-07}	3.72×10^{-11}	1.87×10^{-16}	7.15×10^{-09}	2.95×10^{-11}
		BP	1.42×10^{-07}	1.51×10^{-11}	1.11×10^{-16}	6.09×10^{-10}	3.94×10^{-11}
		CC	4.64×10^{-12}	4.30×10^{-13}	5.30×10^{-17}	4.64×10^{-12}	1.52×10^{-12}
Mouse	$Fmax$	MF	2.58×10^{-04}	4.38×10^{-07}	1.89×10^{-10}	4.09×10^{-03}	4.00×10^{-09}
		BP	1.24×10^{-03}	3.46×10^{-09}	1.16×10^{-13}	1.48×10^{-04}	4.55×10^{-09}
		CC	2.96×10^{-05}	5.33×10^{-07}	1.22×10^{-12}	4.56×10^{-06}	3.73×10^{-07}
	$AUPR$	MF	2.01×10^{-06}	1.51×10^{-10}	2.07×10^{-14}	4.76×10^{-06}	4.98×10^{-11}
		BP	7.61×10^{-04}	5.81×10^{-09}	1.48×10^{-14}	4.25×10^{-06}	5.81×10^{-09}
		CC	1.65×10^{-04}	2.14×10^{-08}	5.98×10^{-15}	1.65×10^{-04}	2.91×10^{-09}
Arabidopsis	$Fmax$	MF	1.74×10^{-08}	5.57×10^{-10}	3.27×10^{-16}	1.76×10^{-09}	1.17×10^{-09}
		BP	2.91×10^{-07}	2.17×10^{-09}	7.00×10^{-12}	1.27×10^{-07}	6.07×10^{-10}
		CC	7.50×10^{-09}	2.30×10^{-10}	1.07×10^{-09}	2.27×10^{-06}	5.81×10^{-10}
	$AUPR$	MF	2.89×10^{-07}	3.47×10^{-11}	2.84×10^{-14}	1.73×10^{-07}	7.94×10^{-11}
		BP	1.76×10^{-06}	6.41×10^{-10}	4.53×10^{-13}	1.37×10^{-07}	1.22×10^{-09}
		CC	1.82×10^{-11}	1.39×10^{-12}	3.84×10^{-15}	6.67×10^{-10}	5.29×10^{-12}
Rat	$Fmax$	MF	3.04×10^{-01}	6.57×10^{-04}	9.63×10^{-07}	1.87×10^{-01}	1.43×10^{-04}
		BP	1.49×10^{-03}	4.76×10^{-05}	1.31×10^{-08}	1.02×10^{-02}	1.09×10^{-05}
		CC	2.04×10^{-01}	2.30×10^{-04}	9.75×10^{-07}	9.68×10^{-03}	2.01×10^{-04}
	$AUPR$	MF	9.02×10^{-04}	3.74×10^{-05}	1.43×10^{-05}	8.78×10^{-06}	1.63×10^{-05}
		BP	1.40×10^{-02}	3.60×10^{-06}	4.04×10^{-10}	1.46×10^{-01}	1.30×10^{-05}
		CC	9.00×10^{-07}	1.44×10^{-10}	1.17×10^{-08}	1.38×10^{-07}	6.65×10^{-11}
Fly	$Fmax$	MF	6.69×10^{-03}	5.35×10^{-06}	9.31×10^{-11}	1.02×10^{-07}	3.97×10^{-07}
		BP	7.87×10^{-06}	6.97×10^{-09}	5.39×10^{-12}	4.91×10^{-10}	1.97×10^{-09}
		CC	3.01×10^{-06}	8.38×10^{-07}	2.77×10^{-14}	1.35×10^{-10}	5.11×10^{-07}
	$AUPR$	MF	3.31×10^{-01}	1.55×10^{-05}	9.90×10^{-14}	6.06×10^{-10}	6.78×10^{-06}
		BP	4.03×10^{-06}	6.66×10^{-10}	4.77×10^{-13}	1.34×10^{-11}	3.24×10^{-10}
		CC	2.04×10^{-08}	2.56×10^{-11}	7.29×10^{-15}	3.58×10^{-13}	3.63×10^{-12}
Budding Yeast	$Fmax$	MF	4.17×10^{-06}	1.05×10^{-08}	5.22×10^{-13}	1.05×10^{-08}	1.22×10^{-08}
		BP	4.32×10^{-06}	7.34×10^{-08}	1.20×10^{-13}	2.07×10^{-08}	1.41×10^{-08}
		CC	3.80×10^{-05}	4.07×10^{-08}	1.41×10^{-11}	1.65×10^{-06}	2.72×10^{-08}
	$AUPR$	MF	9.24×10^{-06}	2.72×10^{-09}	2.75×10^{-15}	1.03×10^{-10}	2.13×10^{-09}
		BP	2.98×10^{-08}	6.08×10^{-11}	2.64×10^{-15}	4.84×10^{-12}	7.64×10^{-11}
		CC	8.16×10^{-05}	4.98×10^{-08}	8.94×10^{-12}	1.13×10^{-08}	2.91×10^{-07}
Fission Yeast	$Fmax$	MF	1.47×10^{-05}	2.70×10^{-02}	4.26×10^{-07}	8.92×10^{-04}	2.70×10^{-02}
		BP	1.05×10^{-06}	4.87×10^{-06}	6.31×10^{-13}	4.48×10^{-08}	4.87×10^{-06}
		CC	9.97×10^{-01}	3.26×10^{-01}	8.52×10^{-04}	3.50×10^{-04}	3.26×10^{-01}
	$AUPR$	MF	6.16×10^{-03}	8.69×10^{-01}	3.91×10^{-08}	8.36×10^{-04}	8.69×10^{-01}
		BP	1.99×10^{-06}	6.24×10^{-06}	4.60×10^{-12}	6.24×10^{-06}	6.24×10^{-06}
		CC	2.41×10^{-03}	2.63×10^{-05}	4.24×10^{-07}	3.51×10^{-06}	2.63×10^{-05}
Nematoda	$Fmax$	MF	5.71×10^{-04}	5.20×10^{-03}	3.19×10^{-08}	2.68×10^{-06}	9.79×10^{-04}
		BP	2.12×10^{-06}	9.91×10^{-07}	1.09×10^{-10}	4.80×10^{-09}	1.58×10^{-07}
		CC	1.45×10^{-02}	1.22×10^{-02}	3.95×10^{-10}	6.02×10^{-07}	2.68×10^{-03}
	$AUPR$	MF	5.09×10^{-04}	8.31×10^{-05}	1.83×10^{-11}	2.19×10^{-08}	1.10×10^{-05}
		BP	2.09×10^{-06}	6.80×10^{-09}	2.96×10^{-13}	5.79×10^{-12}	3.42×10^{-09}
		CC	3.90×10^{-05}	2.28×10^{-07}	6.39×10^{-12}	3.74×10^{-08}	2.28×10^{-07}

Table S3 The p-values between EGPN and other eight GO prediction methods for $WAFmax$ and $WAAUPR$ on the test datasets of 8 species

Measure	GO aspect	(EGPN, EPGP)	(EGPN, GSAGP)	(EGPN, PSAGP)	(EGPN, NGP)	(EGPN, GPN)	(EGPN, EPN)	(EGPN, EGN)	(EGPN, EGP)
$WAFmax$	MF	2.71×10^{-21}	3.85×10^{-18}	2.09×10^{-14}	9.57×10^{-23}	2.53×10^{-10}	7.12×10^{-13}	4.59×10^{-17}	1.70×10^{-04}
	BP	3.39×10^{-15}	4.82×10^{-17}	2.65×10^{-15}	1.42×10^{-18}	9.68×10^{-13}	6.33×10^{-10}	4.37×10^{-12}	4.34×10^{-06}
	CC	1.71×10^{-13}	7.18×10^{-19}	7.49×10^{-18}	2.07×10^{-18}	2.09×10^{-14}	8.98×10^{-10}	5.02×10^{-11}	1.96×10^{-07}
$WAAUPR$	MF	3.31×10^{-24}	3.51×10^{-24}	1.17×10^{-22}	1.16×10^{-25}	3.09×10^{-15}	3.79×10^{-17}	2.31×10^{-20}	3.07×10^{-04}
	BP	6.93×10^{-20}	1.24×10^{-23}	2.99×10^{-22}	2.12×10^{-23}	3.66×10^{-18}	1.52×10^{-14}	4.97×10^{-17}	4.87×10^{-11}
	CC	9.69×10^{-08}	3.81×10^{-16}	3.32×10^{-15}	1.10×10^{-14}	2.70×10^{-10}	1.79×10^{-05}	2.73×10^{-06}	2.77×10^{-04}

Table S4 The p-values between EGPN and other eight GO prediction methods for *Fmax* and *AUPR* on 8 species

Species	Measure	GO aspect	(EGPN, EPGP)	(EGPN, GSAGP)	(EGPN, PSAGP)	(EGPN, NGP)	(EGPN, GPN)	(EGPN, EPN)	(EGPN, EGN)	(EGPN, EGP)
Human	*Fmax*	MF	4.29×10^{-16}	4.45×10^{-14}	4.23×10^{-08}	1.37×10^{-17}	1.58×10^{-03}	1.26×10^{-08}	1.18×10^{-21}	2.65×10^{-03}
		BP	2.04×10^{-15}	1.31×10^{-17}	5.53×10^{-15}	1.75×10^{-18}	7.79×10^{-20}	2.20×10^{-11}	2.76×10^{-19}	5.95×10^{-11}
		CC	9.71×10^{-12}	1.57×10^{-16}	1.58×10^{-15}	8.73×10^{-17}	3.91×10^{-12}	7.89×10^{-09}	9.87×10^{-14}	6.36×10^{-11}
	AUPR	MF	1.18×10^{-21}	2.03×10^{-21}	1.84×10^{-19}	4.08×10^{-23}	9.18×10^{-18}	7.82×10^{-23}	1.05×10^{-31}	3.66×10^{-07}
		BP	2.48×10^{-18}	8.91×10^{-22}	2.70×10^{-20}	1.52×10^{-21}	1.18×10^{-26}	1.42×10^{-17}	8.86×10^{-25}	1.68×10^{-15}
		CC	5.75×10^{-04}	3.72×10^{-12}	1.71×10^{-11}	3.28×10^{-11}	5.69×10^{-08}	4.07×10^{-03}	6.30×10^{-03}	5.40×10^{-01}
Mouse	*Fmax*	MF	5.63×10^{-21}	3.45×10^{-15}	8.18×10^{-12}	3.72×10^{-22}	1.36×10^{-04}	2.08×10^{-17}	3.99×10^{-23}	1.42×10^{-02}
		BP	3.56×10^{-16}	2.99×10^{-16}	1.96×10^{-16}	6.38×10^{-19}	1.74×10^{-12}	1.26×10^{-12}	3.48×10^{-16}	6.33×10^{-06}
		CC	1.06×10^{-12}	1.69×10^{-16}	8.19×10^{-17}	1.79×10^{-17}	2.70×10^{-18}	1.56×10^{-12}	1.98×10^{-10}	6.37×10^{-01}
	AUPR	MF	3.98×10^{-22}	2.04×10^{-20}	1.73×10^{-19}	4.01×10^{-23}	2.16×10^{-09}	1.19×10^{-25}	5.99×10^{-28}	3.62×10^{-01}
		BP	2.60×10^{-19}	9.96×10^{-22}	1.70×10^{-21}	2.96×10^{-22}	3.74×10^{-22}	1.12×10^{-22}	2.83×10^{-23}	3.21×10^{-11}
		CC	7.71×10^{-06}	9.57×10^{-14}	2.25×10^{-13}	3.96×10^{-13}	1.28×10^{-11}	1.43×10^{-03}	2.78×10^{-03}	2.79×10^{-01}
Arabidopsis	*Fmax*	MF	1.52×10^{-17}	5.58×10^{-14}	1.06×10^{-12}	4.93×10^{-19}	1.39×10^{-11}	7.77×10^{-15}	7.09×10^{-17}	2.28×10^{-03}
		BP	3.80×10^{-11}	1.67×10^{-10}	5.28×10^{-11}	1.01×10^{-14}	7.19×10^{-09}	1.04×10^{-07}	1.39×10^{-07}	8.80×10^{-02}
		CC	1.58×10^{-07}	5.08×10^{-14}	1.80×10^{-14}	1.02×10^{-12}	3.19×10^{-13}	5.82×10^{-08}	2.30×10^{-03}	4.44×10^{-02}
	AUPR	MF	1.07×10^{-20}	1.21×10^{-21}	2.38×10^{-20}	1.78×10^{-22}	1.15×10^{-14}	1.19×10^{-22}	1.41×10^{-26}	3.21×10^{-06}
		BP	3.96×10^{-18}	3.10×10^{-21}	3.17×10^{-20}	1.39×10^{-21}	1.62×10^{-24}	2.99×10^{-22}	5.88×10^{-22}	1.49×10^{-13}
		CC	7.76×10^{-05}	3.35×10^{-14}	2.98×10^{-13}	1.41×10^{-11}	1.52×10^{-11}	3.53×10^{-01}	2.32×10^{-03}	1.20×10^{-01}
Rat	*Fmax*	MF	8.51×10^{-18}	3.25×10^{-11}	1.73×10^{-11}	4.62×10^{-19}	8.20×10^{-01}	6.32×10^{-15}	7.91×10^{-16}	4.98×10^{-01}
		BP	5.04×10^{-10}	2.14×10^{-11}	4.67×10^{-11}	4.26×10^{-14}	6.72×10^{-08}	2.01×10^{-06}	3.11×10^{-07}	2.46×10^{-01}
		CC	2.12×10^{-08}	3.22×10^{-11}	6.12×10^{-12}	2.22×10^{-12}	1.50×10^{-08}	1.23×10^{-03}	1.24×10^{-02}	8.58×10^{-04}
	AUPR	MF	1.22×10^{-22}	6.41×10^{-21}	6.20×10^{-20}	1.43×10^{-23}	4.55×10^{-09}	2.09×10^{-28}	9.71×10^{-27}	2.84×10^{-09}
		BP	1.03×10^{-16}	5.61×10^{-20}	3.12×10^{-19}	3.76×10^{-20}	1.05×10^{-21}	5.15×10^{-14}	5.67×10^{-20}	1.52×10^{-03}
		CC	5.31×10^{-09}	2.13×10^{-15}	2.28×10^{-15}	6.98×10^{-15}	1.08×10^{-14}	5.16×10^{-07}	1.91×10^{-02}	5.95×10^{-02}
Fly	*Fmax*	MF	1.60×10^{-18}	1.42×10^{-15}	1.19×10^{-12}	1.83×10^{-20}	6.27×10^{-12}	5.87×10^{-12}	5.21×10^{-19}	7.79×10^{-06}
		BP	2.13×10^{-12}	5.73×10^{-15}	9.23×10^{-13}	2.09×10^{-16}	1.57×10^{-14}	6.39×10^{-09}	4.25×10^{-16}	4.74×10^{-05}
		CC	1.00×10^{-09}	5.94×10^{-15}	4.35×10^{-12}	1.00×10^{-14}	3.24×10^{-16}	8.24×10^{-03}	8.18×10^{-15}	1.64×10^{-05}
	AUPR	MF	3.30×10^{-22}	2.71×10^{-22}	3.56×10^{-21}	4.08×10^{-24}	6.16×10^{-26}	5.94×10^{-25}	5.07×10^{-25}	9.30×10^{-11}
		BP	2.32×10^{-15}	1.09×10^{-19}	5.40×10^{-18}	1.61×10^{-19}	1.61×10^{-25}	2.45×10^{-10}	1.16×10^{-19}	3.01×10^{-12}
		CC	2.59×10^{-10}	3.70×10^{-19}	1.35×10^{-17}	6.36×10^{-18}	9.80×10^{-21}	4.82×10^{-02}	1.16×10^{-15}	1.02×10^{-04}
Budding Yeast	*Fmax*	MF	3.34×10^{-16}	4.96×10^{-16}	1.04×10^{-11}	4.28×10^{-18}	5.01×10^{-15}	5.42×10^{-09}	3.43×10^{-20}	8.27×10^{-04}
		BP	8.17×10^{-10}	1.10×10^{-14}	1.31×10^{-11}	3.14×10^{-14}	1.76×10^{-14}	7.76×10^{-01}	1.91×10^{-11}	4.07×10^{-02}
		CC	8.42×10^{-02}	2.97×10^{-12}	1.29×10^{-07}	1.34×10^{-09}	1.41×10^{-08}	9.25×10^{-01}	1.59×10^{-02}	8.00×10^{-03}
	AUPR	MF	1.15×10^{-17}	3.63×10^{-20}	1.53×10^{-18}	4.30×10^{-20}	1.61×10^{-22}	1.23×10^{-15}	1.43×10^{-25}	5.33×10^{-01}
		BP	6.37×10^{-15}	4.23×10^{-21}	4.02×10^{-19}	8.61×10^{-20}	2.08×10^{-26}	9.02×10^{-11}	1.77×10^{-21}	1.23×10^{-09}
		CC	5.34×10^{-17}	3.81×10^{-25}	1.98×10^{-23}	2.54×10^{-23}	8.53×10^{-22}	1.22×10^{-05}	2.16×10^{-23}	4.61×10^{-14}
Fission Yeast	*Fmax*	MF	5.83×10^{-20}	1.68×10^{-16}	8.72×10^{-01}	1.95×10^{-20}	6.87×10^{-03}	8.26×10^{-13}	1.14×10^{-25}	6.42×10^{-01}
		BP	3.00×10^{-12}	2.56×10^{-14}	2.44×10^{-10}	3.25×10^{-14}	1.66×10^{-07}	2.30×10^{-04}	3.11×10^{-18}	1.88×10^{-02}
		CC	5.15×10^{-13}	1.71×10^{-16}	3.83×10^{-13}	3.41×10^{-15}	3.03×10^{-12}	4.83×10^{-05}	7.40×10^{-16}	3.99×10^{-11}
	AUPR	MF	6.64×10^{-20}	4.96×10^{-20}	4.11×10^{-18}	1.21×10^{-20}	4.03×10^{-09}	1.86×10^{-16}	2.11×10^{-28}	1.63×10^{-07}
		BP	1.93×10^{-18}	3.65×10^{-22}	2.38×10^{-19}	5.18×10^{-21}	7.43×10^{-17}	1.56×10^{-11}	3.91×10^{-27}	1.41×10^{-08}
		CC	5.39×10^{-14}	1.01×10^{-19}	9.20×10^{-18}	2.48×10^{-18}	9.57×10^{-22}	3.94×10^{-01}	1.37×10^{-18}	5.73×10^{-07}
Nematoda	*Fmax*	MF	3.62×10^{-14}	6.79×10^{-11}	8.72×10^{-01}	8.85×10^{-16}	8.91×10^{-01}	7.73×10^{-08}	6.60×10^{-18}	7.18×10^{-03}
		BP	2.31×10^{-08}	9.59×10^{-12}	9.80×10^{-10}	1.39×10^{-12}	4.25×10^{-12}	5.61×10^{-03}	1.40×10^{-09}	7.70×10^{-01}
		CC	8.47×10^{-08}	6.57×10^{-15}	4.39×10^{-12}	9.90×10^{-13}	3.53×10^{-07}	1.69×10^{-02}	1.04×10^{-10}	1.90×10^{-03}
	AUPR	MF	2.34×10^{-17}	1.28×10^{-17}	2.23×10^{-13}	1.34×10^{-18}	1.00×10^{-01}	2.24×10^{-01}	3.21×10^{-23}	5.07×10^{-02}
		BP	6.74×10^{-14}	5.92×10^{-19}	4.85×10^{-17}	3.38×10^{-18}	3.54×10^{-22}	6.00×10^{-08}	1.48×10^{-17}	5.85×10^{-02}
		CC	1.33×10^{-09}	1.28×10^{-18}	3.85×10^{-16}	3.71×10^{-16}	2.44×10^{-15}	5.88×10^{-01}	1.81×10^{-10}	9.70×10^{-02}

Table S5 The p-values between EGN and other six GO prediction methods for $Fmax$ and $AUPR$ on 98 non-coding genes

Measure	GO aspect	(EGN, EPGP)	(EGN, GSAGP)	(EGN, NGP)	(EGN, GN)	(EGN, EN)	(EGN, EG)
$WAFmax$	MF	1.09×10^{-12}	4.01×10^{-17}	5.48×10^{-18}	4.01×10^{-17}	1.95×10^{-12}	4.16×10^{-10}
	BP	3.86×10^{-10}	3.36×10^{-18}	1.07×10^{-16}	7.98×10^{-17}	3.86×10^{-10}	1.39×10^{-03}
	CC	1.73×10^{-05}	1.35×10^{-13}	1.35×10^{-13}	1.17×10^{-11}	1.80×10^{-03}	2.80×10^{-05}
$WAAUPR$	MF	4.83×10^{-20}	3.39×10^{-25}	2.87×10^{-25}	4.30×10^{-24}	1.13×10^{-19}	6.30×10^{-17}
	BP	5.90×10^{-09}	1.51×10^{-23}	1.99×10^{-21}	9.56×10^{-21}	1.38×10^{-09}	5.88×10^{-11}
	CC	1.56×10^{-11}	8.53×10^{-25}	2.93×10^{-24}	3.07×10^{-22}	1.56×10^{-11}	1.01×10^{-05}

Table S6 The numbers of genes with GO annotation of three aspects for 20 species

Database	Species	Version	Gene number	Sample number	GO number	MF number	BP number	CC number
COXPRESdb	Nematoda	Cel-m.c4-0	17256	1780	3154	1254	2705	2018
	Dog	Cfa-m.c3-0	16214	777	96	31	54	79
	Fly	Dme-m.c4-0	12626	4209	5317	2729	4874	3495
	Zebrafish	Dre-m.c4-0	10112	1423	2477	541	2324	414
	Chicken	Gga-m.c4-0	13757	1502	502	215	383	337
	Human	Hsa-m2.c3-0	20199	27655	14706	9281	12362	13278
	Monkey	Mcc-m.c3-0	15782	1006	0	0	0	0
	Mouse	Mmu-m.c4-0	20962	42916	10564	5646	8909	7621
	Rat	Rno-m.c4-0	13751	42752	5409	3594	4387	4135
	Budding yeast	Sce-m.c3-0	4461	3593	4107	3130	3934	3402
	Fission yeast	Spo-m.c3-0	4881	166	2743	1303	2339	1877
ATTED-II	Arabidopsis	Ath-m.c8-0	20819	12686	11602	5090	7927	8656
	Field mustard	Bra-r.c3-0	26339	164	0	0	0	0
	Soybean	Gma-m.c4-0	15746	1022	0	0	0	0
	Medicago	Mtr-m.c4-1	20376	780	0	0	0	0
	Rice	Osa-m.c7-0	19867	1775	82	59	69	55
	Poplar	Ppo-m.c3-0	21910	557	0	0	0	0
	Tomato	Sly-m.c4-0	5721	392	0	0	0	0
	Grape	Vvi-m.c4-0	9421	258	0	0	0	0
	Maize	Zma-m.c4-0	10777	606	0	0	0	0

Gene number: the total number of genes in a species.

Sample number: the number of experimental samples in microarray technology.

GO number: the number of genes with GO annotation in a species.

MF/BP/CC number: the number of genes with MF/BP/CC GO annotation in a species.

Table S7 The details of 8 benchmark datasets

Species	NTR[1]	NEV[1]	NTE[1]	NGT_MF[2]	NGT_BP[3]	NGT_CC[4]
Human	12501	735	1470	3841	11674	1505
Mouse	8965	527	1054	2735	12035	1188
Arabidopsis	9862	580	1160	2245	4787	563
Rat	4599	270	540	2542	7917	954
Fly	4521	265	531	1783	6123	857
Budding Yeast	3492	205	410	2025	4525	899
Fission Yeast	2332	137	274	1426	3885	720
Nematoda	2682	157	315	1160	4042	539

[1]NTR/NEV/NTE: the number of genes in training/validation/test datasets

[2]NGT_MF: the total number of MF terms in training, validation and test datasets

[3]NGT_BP: the total number of BP terms in training, validation and test datasets

[4]NGT_CC: the total number of CC terms in training, validation and test datasets

Table S8 The values of α, h, $margin$, and c_f on the benchmark datasets for 8 species

Species	GO aspect	α	h	$margin$	c_f
Human	MF	5	1000	0.01	0.90
	BP	5	1000	0.01	0.90
	CC	10	1000	0.01	0.95
Mouse	MF	5	1000	0.01	0.90
	BP	3	1000	0.01	0.90
	CC	2	1000	0.01	0.95
Arabidopsis	MF	5	1000	0.01	0.90
	BP	5	1000	0.01	0.90
	CC	5	1000	0.01	0.95
Rat	MF	3	1000	0.01	0.90
	BP	5	1000	0.01	0.90
	CC	5	1000	0.01	0.95
Fly	MF	3	1000	0.01	0.90
	BP	5	1000	0.01	0.90
	CC	5	1000	0.01	0.95
Budding Yeast	MF	3	1000	0.01	0.90
	BP	5	1000	0.01	0.90
	CC	5	1000	0.01	0.95
Fission Yeast	MF	3	-	0.01	0.90
	BP	5	-	0.01	0.90
	CC	5	-	0.01	0.95
Nematoda	MF	5	1000	0.01	0.90
	BP	5	1000	0.01	0.90
	CC	5	1000	0.01	0.95

'-' means that the PCA is not executed in the corresponding species

Supporting Figures

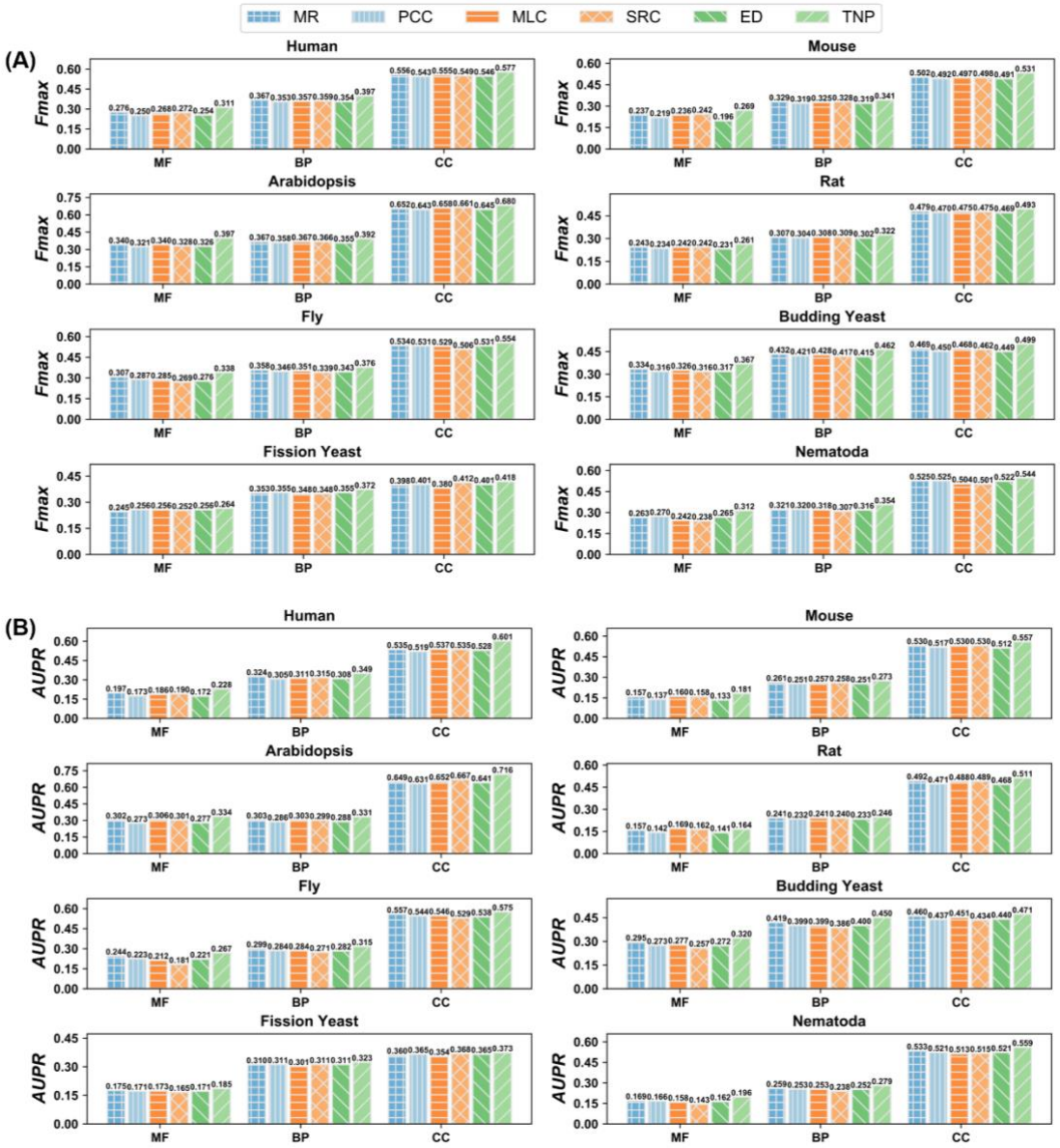

Figure S1　The performance of six expression profile-based methods on the test datasets for 8 species

A. The $Fmax$ values of six methods for 8 species. **B**. The $AUPR$ values of six methods for 8 species.

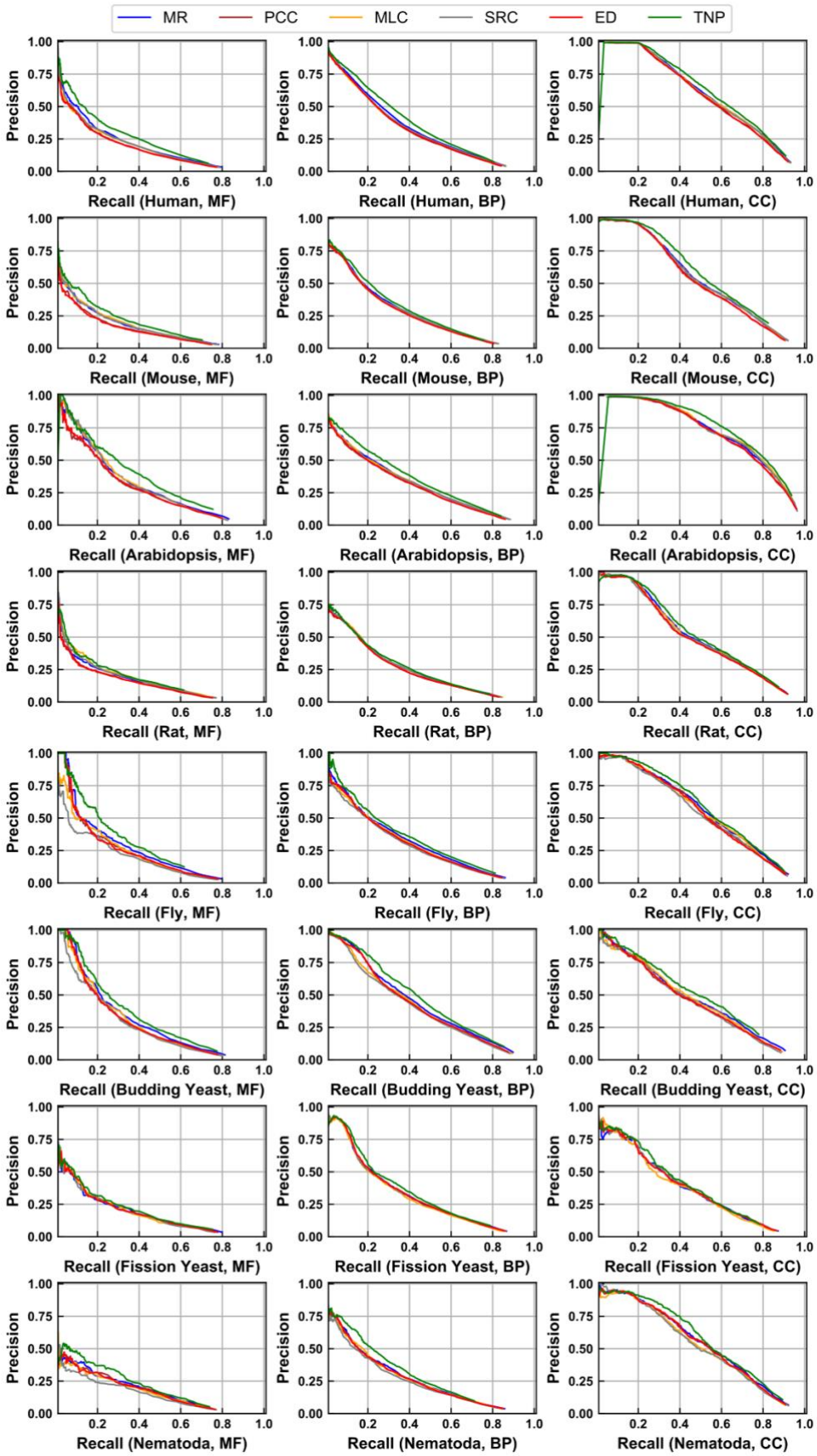

Figure S2 The precision-recall curves of six expression profile-based methods on the test datasets for 8 species

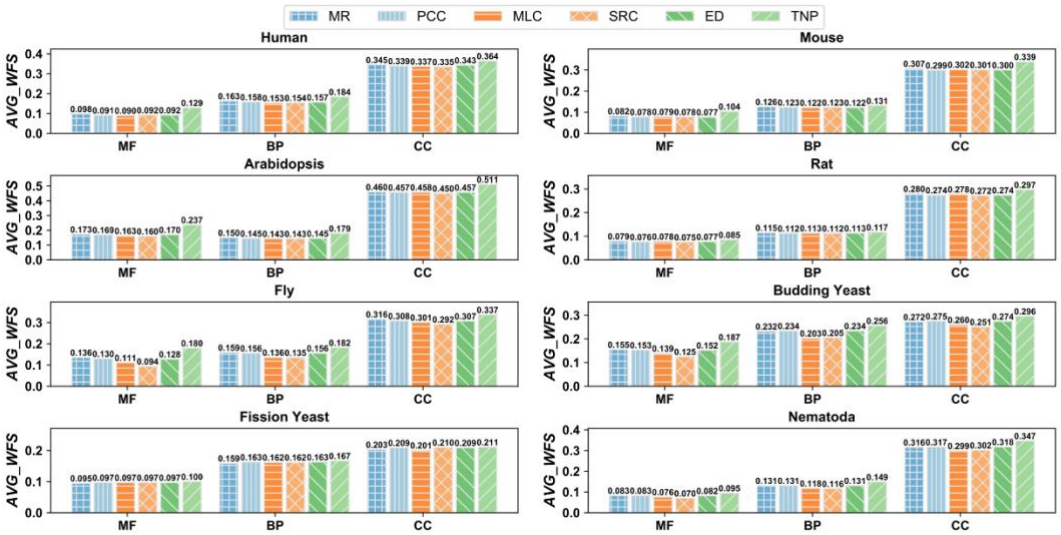

Figure S3 The *AVG_WFS* values of six measures for three GO aspects in 8 individual species

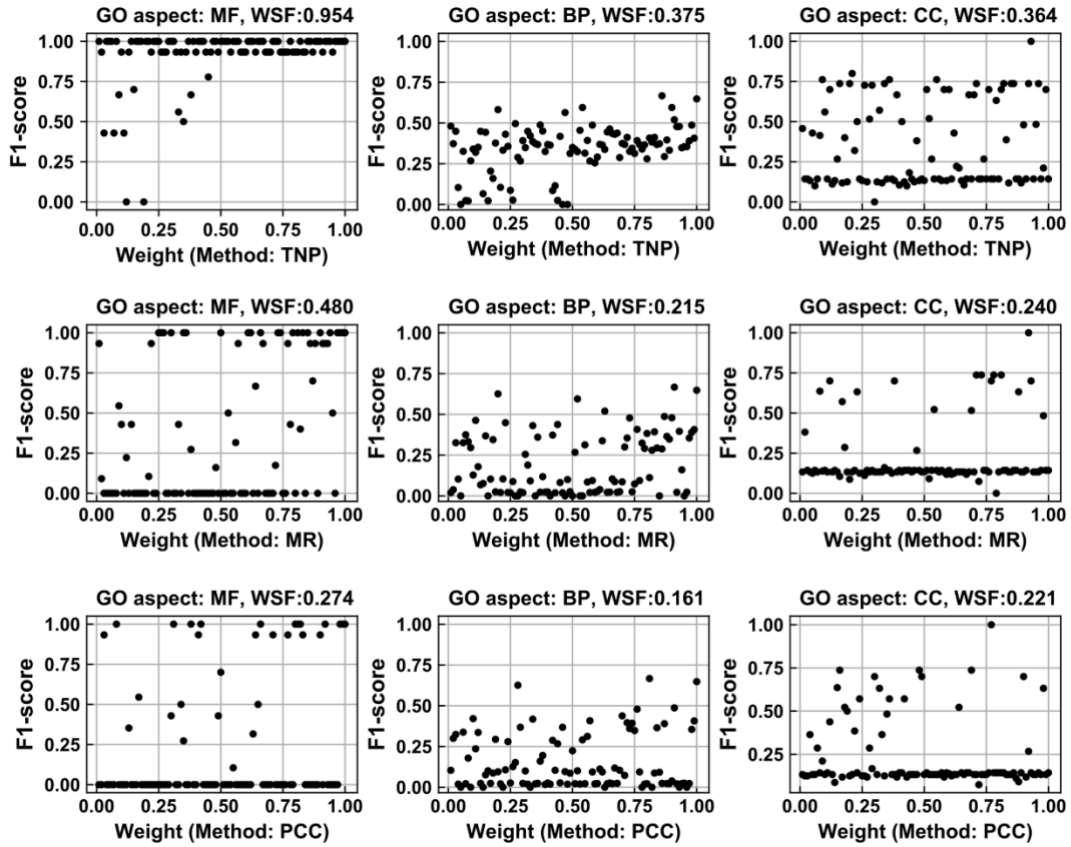

Figure S4 The scattering plots of weights versus F1-scores of 100 templates for the gene MIRLET7C over TNP, MR, and PCC

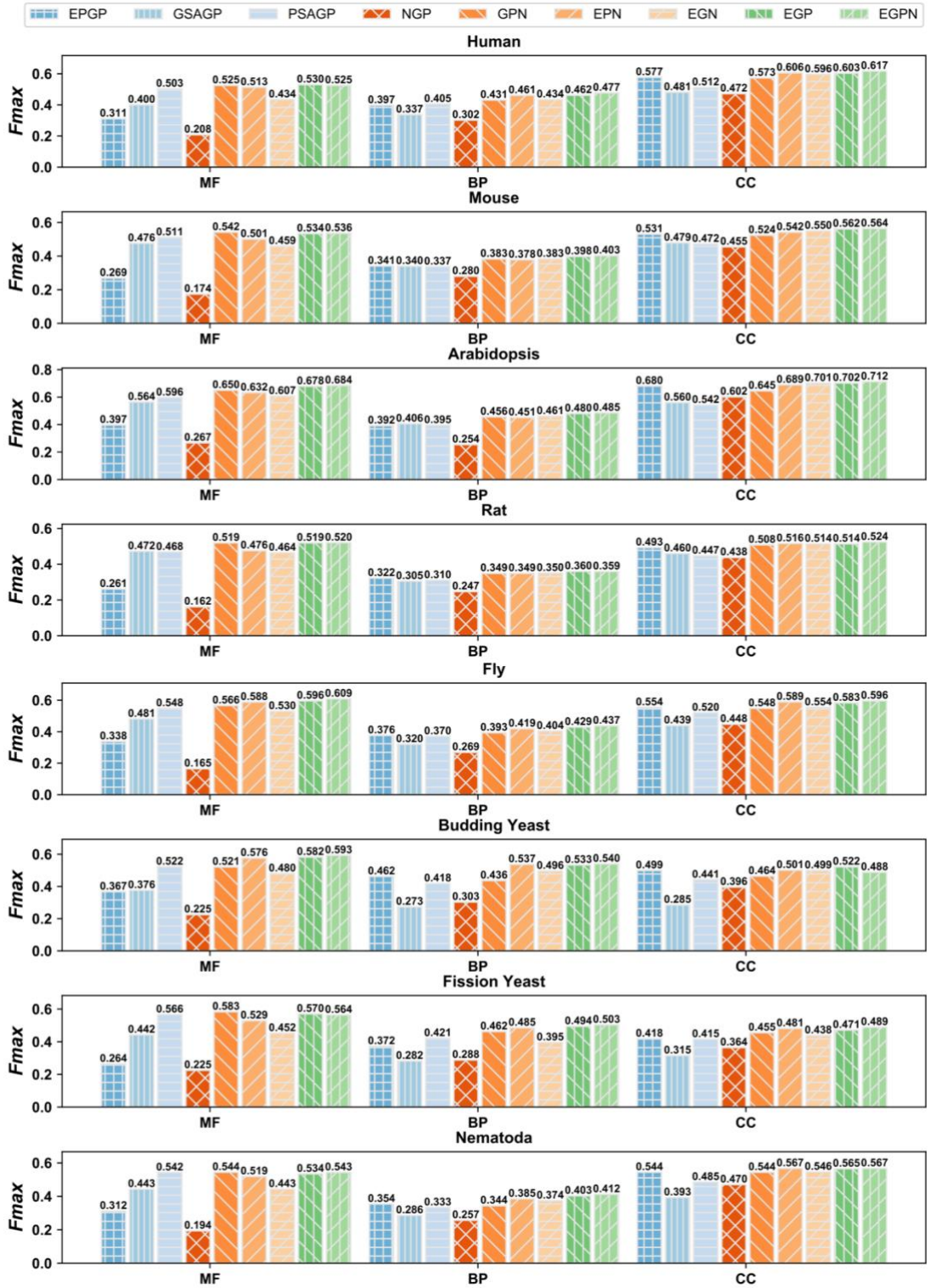

Figure S5 The $Fmax$ values of nine GO prediction methods on the test datasets for 8 species

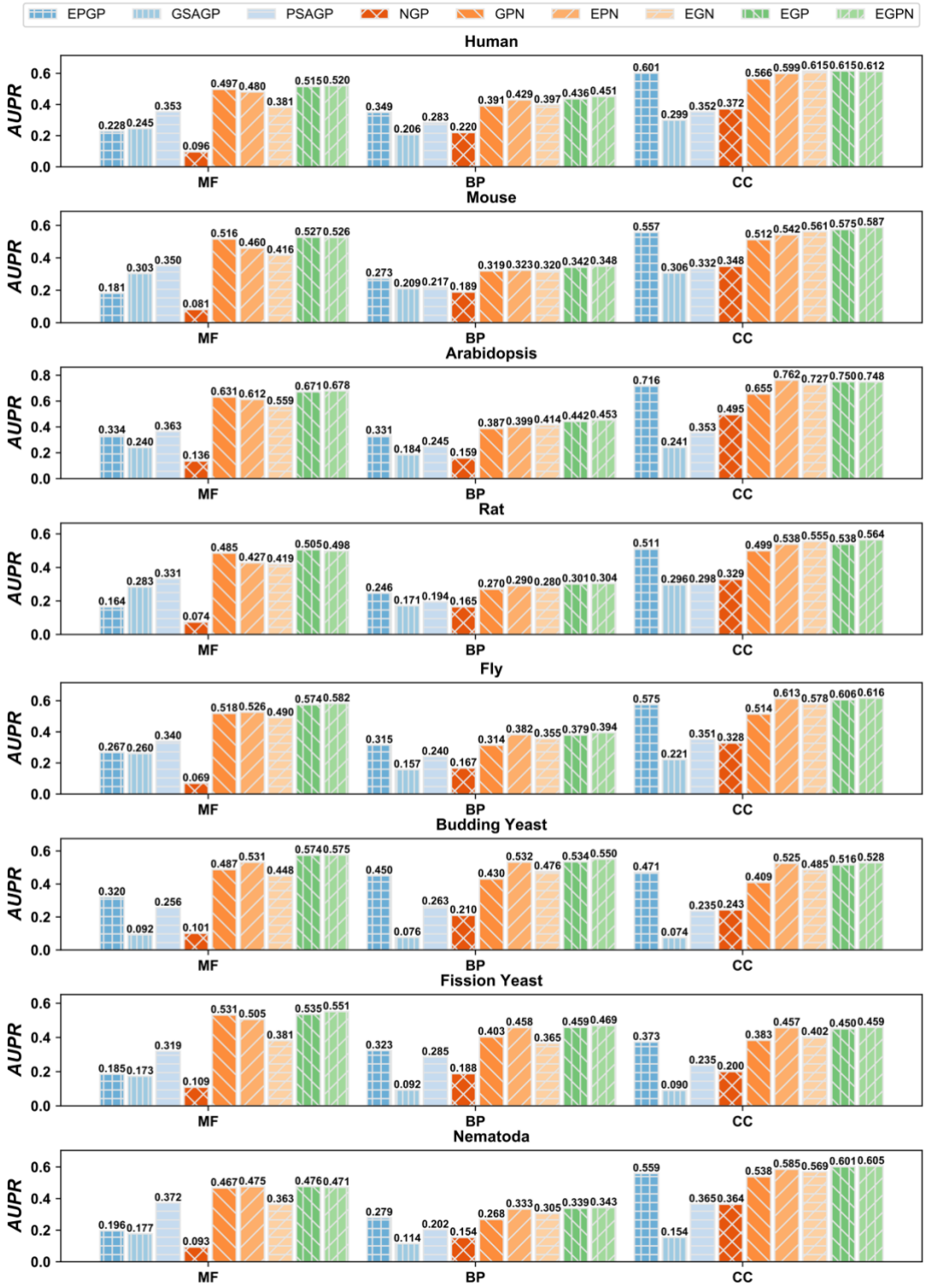

Figure S6 The *AUPR* values of nine GO prediction methods on the test datasets for 8 species

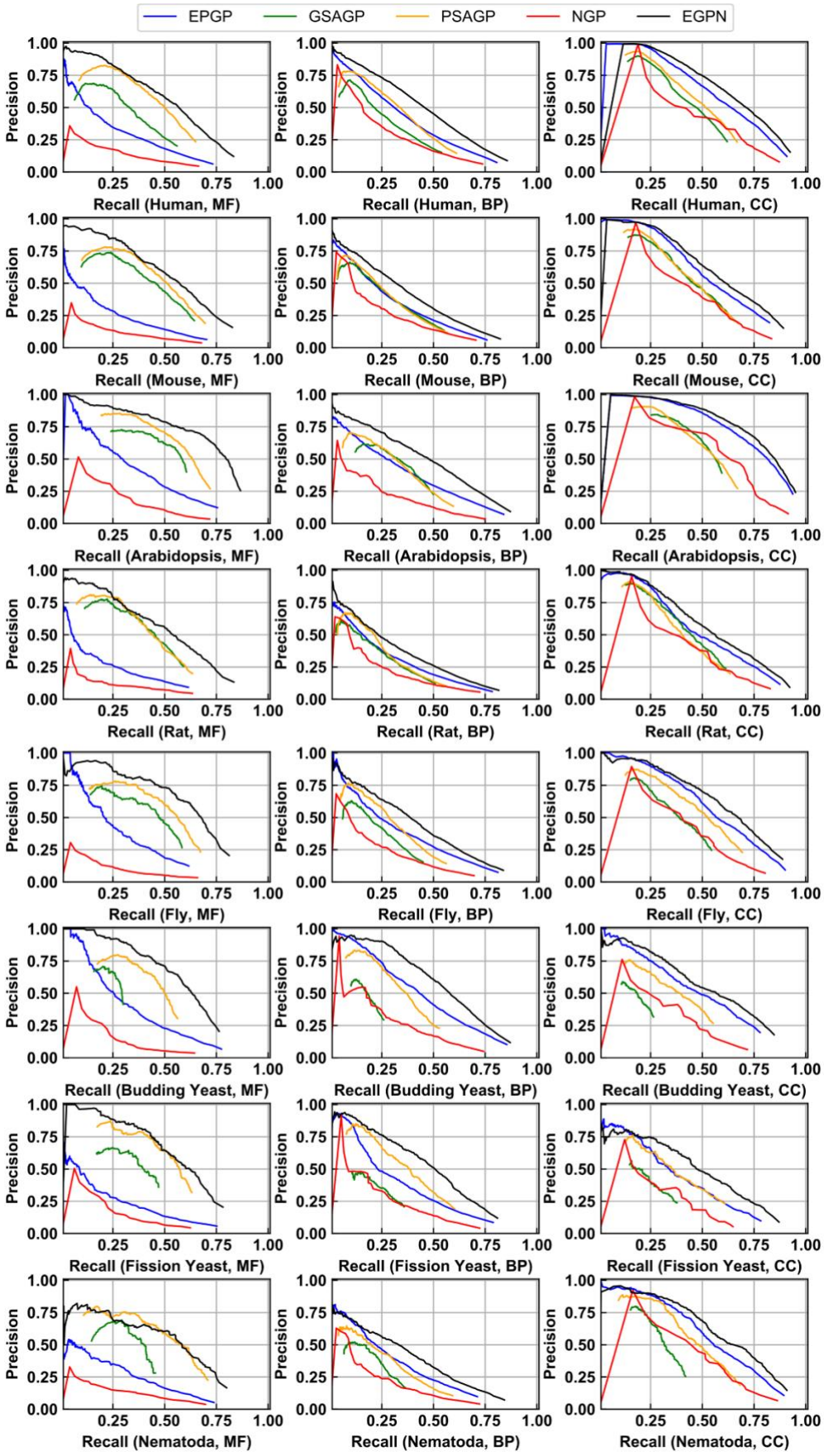

Figure S7 The precision-recall curves of five GO prediction methods on the test datasets for 8 species.

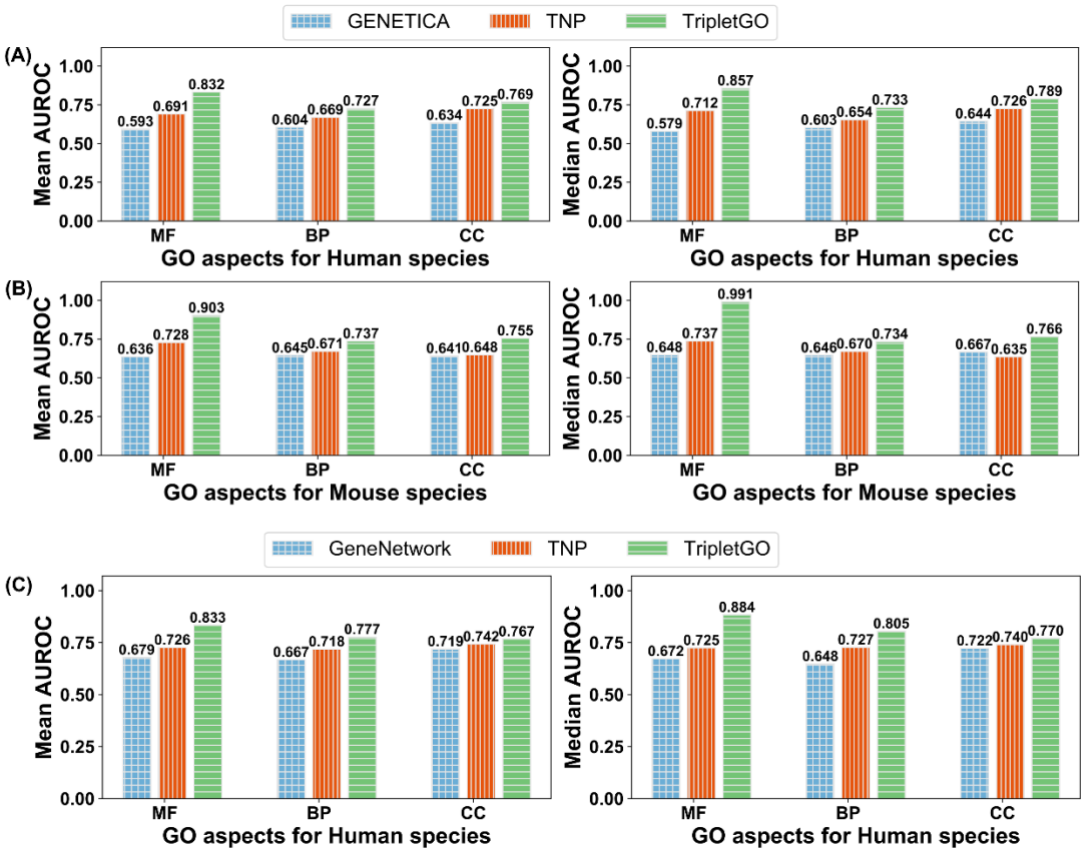

Figure S8 Comparison of mean and median AUROC values of three GO aspects by different methods on the common dataset

A. GENETICA, TNP and TripletGO on human; **B**. GENETICA, TNP and TripletGO on mouse; **C**. GeneNetwork, TNP and TripletGO on human.

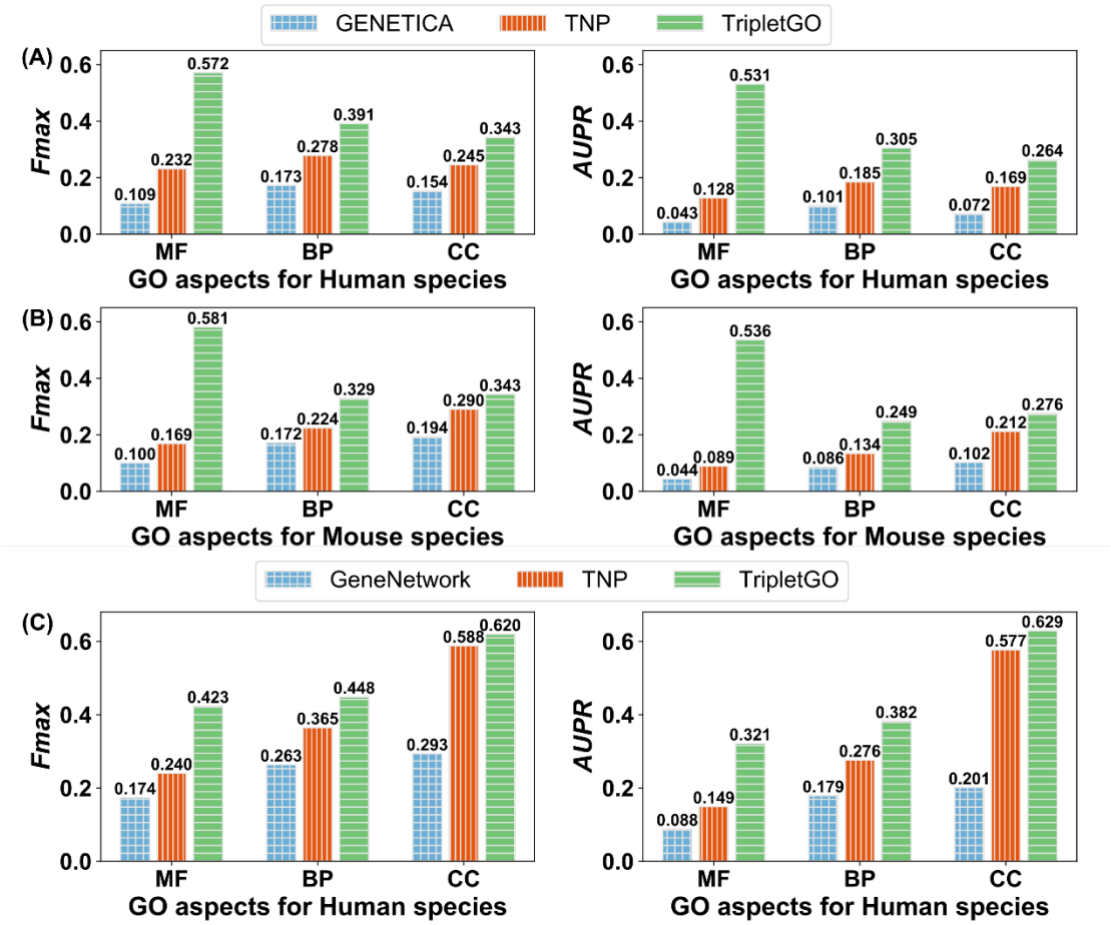

Figure S9 Comparison of *Fmax* and *AUPR* values of three GO aspects by different methods on the common dataset

A. GENETICA, TNP and TripletGO on human; **B**. GENETICA, TNP and TripletGO on mouse; **C**. GeneNetwork, TNP and TripletGO on human.

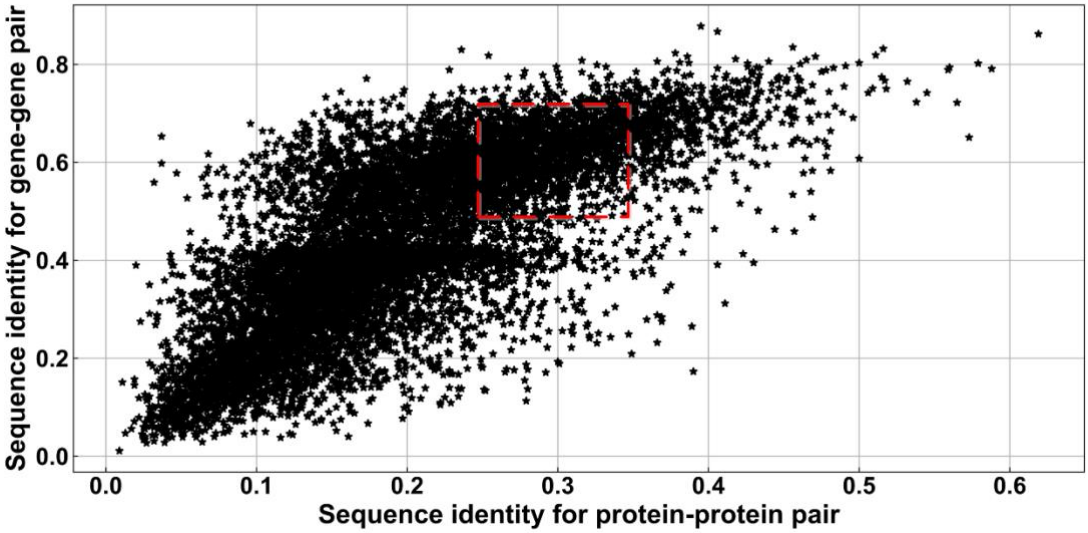

Figure S10 The distribution of sequence identities for 10000 gene-gene pairs and 10000 mapped protein-protein pairs

Supporting References

[1] Heller MJ. DNA microarray technology: devices, systems, and applications. Annu Rev Biomed Eng 2002; 4: 129-53.

[2] Patro S, Sahu KK. Normalization: A preprocessing stage. arXiv preprint arXiv:1503.06462 2015.

[3] Wold S, Esbensen K, Geladi P. Principal component analysis. Chemometr Intell Lab Syst 1987; 2: 37-52.

[4] Adler J, Parmryd I. Quantifying colocalization by correlation: the Pearson correlation coefficient is superior to the Mander's overlap coefficient. Cytometry A 2010; 77: 733-42.

[5] Zar JH. Significance testing of the Spearman rank correlation coefficient. J Am Stat Assoc 1972; 67: 578-80.

[6] Obayashi T, Aoki Y, Tadaka S, Kagaya Y, Kinoshita K. ATTED-II in 2018: a plant coexpression database based on investigation of the statistical property of the mutual rank index. Plant Cell Physiol 2018; 59: e3-e.

[7] Wang L, Zhang Y, Feng J. On the Euclidean distance of images. IEEE Trans Pattern Anal Mach Intell 2005; 27: 1334-9.

[8] Makrodimitris S, Reinders MJ, van Ham RC. Metric learning on expression data for gene function prediction. Bioinformatics 2020; 36: 1182-90.

[9] Obayashi T, Kagaya Y, Aoki Y, Tadaka S, Kinoshita K. COXPRESdb v7: a gene coexpression database for 11 animal species supported by 23 coexpression platforms for technical evaluation and evolutionary inference. Nucleic Acids Res 2019; 47: D55-D62.

[10] Byrd RH, Lu P, Nocedal J, Zhu C. A limited memory algorithm for bound constrained optimization. SIAM J Sci Comput 1995; 16: 1190-208.

[11] Ruxton GD. The unequal variance t-test is an underused alternative to Student's t-test and the Mann–Whitney U test. Behav Ecol 2006; 17: 688-90.

[12] Heeren T, D'Agostino R. Robustness of the two independent samples t-test when applied to ordinal scaled data. Stat Med 1987; 6: 79-90.

[13] Crawford J, Howell DC, Garthwaite PH. Payne and Jones revisited: estimating the abnormality of test score differences using a modified paired samples t test. J Clin Exp Neuropsychol 1998; 20: 898-905.

[14] Sayers EW, Barrett T, Benson DA, Bolton E, Bryant SH, Canese K, et al. Database resources of the national center for biotechnology information. Nucleic Acids Res 2010; 39: D38-D51.

[15] Radivojac P, Clark WT, Oron TR, Schnoes AM, Wittkop T, Sokolov A, et al. A large-scale evaluation of computational protein function prediction. Nat Methods 2013; 10: 221-7.

[16] Jiang Y, Oron TR, Clark WT, Bankapur AR, D'Andrea D, Lepore R, et al. An expanded evaluation of protein function prediction methods shows an improvement in accuracy. Genome Biol 2016; 17: 1-19.

[17] Consortium GO. The Gene Ontology (GO) database and informatics resource. Nucleic Acids Res 2004; 32: D258-D61.

[18] Consortium U. UniProt: a hub for protein information. Nucleic Acids Res 2015; 43: D204-D12.

[19] Montgomery DC, Peck EA, Vining GG. Introduction to linear regression analysis2021.

[20] Awad M, Khanna R. Support vector regression. Efficient learning machines 2015: 67-80.

[21] Needleman SB, Wunsch CD. A general method applicable to the search for similarities in the amino acid sequence of two proteins. J Mol Biol 1970; 48: 443-53.