

Each transporter gene might be associated with several TC family (or sub-family) numbers, each with a different score. Therefore, a method for the assignment of TC family numbers, with the intention of expediting the identification of genes associated to specific functions, based on the computation of the weighted score for each family is proposed and calculated as:

$$\text{weighted score (\%)} = \frac{\text{score}}{\sum_{i=0}^n \text{score}_i} \times 100.$$

The denominator is the sum of all the TC family number scores for a given gene. If the weighted score of the highest TC family number is less than 50%, the protein is annotated with more than one TC family number. That annotation is achieved by adding subsequent TC family numbers in the list of highest scores, until the cumulative sum surpasses 50%. As an example, if a TCG has similarities to carriers classified in 3 distinct TC families: 2.A.16.4.# with a score of 28%; 2.A.16.2.# with 39% and 2.A.16.3.# with 33%, there is no TC family with a score of more than 50%. Thus, the protein is annotated with the two highest TC families since the sum of the scores is over 50% (72%). It is important to stress that the metabolites' scoring process is independent of the TC family score. The TC number assigned to the TCG presented in Table 1A and Table 2A of the supplemental material is 2.A.1.1.# with a score of 100%.